

Environmental policy still undefined

The British Government, which has several problems on its mind, does not need new environmental legislation. Its Secretary of State for the Environment has probably ensured that the embarrassment can be postponed.

POOR Mr Christopher Patten, British Secretary of State for the Environment and the darling of British environmental lobbyists since his speech at the government party's conference a year ago, has unjustly been given a bad press for last week's white paper on environmental policy (see page 412). More accurately, he has been given a bad press for the wrong reasons. Patten's fickle lobbyist friends, as ever zealous in their pursuit of their ideal and tightly regulated world, are demanding to know why he has not committed the British government to such measures as the carbon tax they have been pushing at him for the past year, forgetting that Mr Saddam Hussein may have done that instead.

The more serious complaint, quite buried in the chorus of mistaken umbrage, is that the white paper fails to provide a framework within which this or any future British government can determine environmental policy in a coherent fashion. For a document subtitled "Britain's Environmental Strategy" (the title proper, *This Common Inheritance*, is taken from Stuart Mill), that is at best an oversight.

The principles, sadly, are not principles at all, but platitudes (yet nicely written). Almost the most daring statement in the document, but one that may sufficiently explain the disaffection of the lobbyists from Patten's way of seeing things, is that economic growth is taken to engender popular concern for environmental protection. Or, to be fair both to Patten and the lobbyists, as "people have enjoyed the benefits of economic growth, they have become increasingly preoccupied with the quality of their daily lives". From there, it is but a short step, by way of uplifting but cramping illustrative examples ("aspirations" for healthier living, "desire" for cleaner air, "enthusiasm" for protecting the best of "our" surroundings, "hope" that we can pass on the best of what we value most to our children), to a thoroughly middle-class view of the environmental problem.

What about the fellow (man or woman) who has it in his head that the benefits of economic growth are most aptly typified by the speed at which an affordable motorcar will travel from A to B? Or the others who will not cross the Atlantic except supersonically? Will aspirations to "healthier living" and the rest win them away from their inclinations? Of course not. And that is why, for much of the world, environmental strategy is what the middle classes worry about because they have been freed from

the preoccupation with survival. (Immortality will be next on the agenda.)

The serious disappointment is what the white paper has to say, or has decided not to say, about strategic issues. If, at the most primitive level, environmental protection is about personal survival, the age-specific likelihood of death must be a useful index. So why not work out, using empirical as well as theoretical models, the chances that people will die or be incapacitated by this or that environmental insult? Is ubiquitous radioactivity more damaging than, say, ubiquitous PCBs? And by what factor? There is no compelling reason why a government should have an opinion on such an issue, but at least it might create a mechanism for resolving it. Apparently there is none.

Worse still, while *This Common Inheritance* glossily gives prominence to the problem of global warming, which may (or may not) be a consequence of the increasing concentration of carbon dioxide in the atmosphere, and while it properly acknowledges that the small chance of being wrong on such a huge issue could have consequences that would be calamitous, it does very little to help those who will read it to know where to put their money, or their energy and intelligence. Might it not have been politically as well as philosophically more prudent to have come to a decision and to have said that global warming is more (or less) important than clean streets, or one of the other 300 talismans of environmental probity in this policy-free policy statement?

Only against this background of lost opportunity does the white paper's strongest suit shine through, but the document is strong on the preservation of the appearance of the country over which the British government presides. Indeed, in that respect, even the philosophy makes sense — "the reason for keeping it like it is consists of people's inclinations", which is a kind of democratic principle. But in this connection, the white paper has very little to say about the anomalies of the British planning system (meant to prevent unpopular development), the associated economic issues (such as how a landowner prevented from development should be compensated, and by whom) and trivial but teasing questions such as how much effort should be spent in preserving British landscapes and appearances when they are replicated many times in the Europe to which Britain will soon belong. □



UN and the future

The present halcyon state of UN members' minds is an opportunity for administrative reform.

PREVAILING cheerfulness at the United Nations about the almost solid front against Iraq during last week's opening speeches to the annual General Assembly should not blind the rest of us to the defects of the organization that comes closest to the concept of world government. At best, it is an historical accident that most members of the United Nations agree that Iraq's annexation of Kuwait is an offence against international law. It is too soon to assume, as UN well-wishers do, that good sense will now prevail for the rest of time. Indeed, the Soviet Union which has made consensus possible in New York may not be about, in its present shape, when the next invasion happens.

That is why well-wishers have one most urgent need — to make the institutions of the United Nations effective and efficient. It is not unremarkable that most of these are technical agencies. The oldest is the World Meteorological Office, brought into being while Bismark was still in power by the felt need to share data about the weather. On balance, it is a sensible outfit, even if it has strayed from the path of good sense during its recent intervention in the global warming problem. The larger technical organizations, the World Health Organization and the Food and Agriculture Organization, do some good works, but many fewer than they might with all that money. Unesco, despite the efforts of the most able secretary-general ever (which is not very long), remains a dead end. While the politicians are happy with events in the Middle East, and therefore compliant, might not this be the time to embark on the radical reappraisal of what the specialized agencies are for, and how they should operate, that the circumstances have required for several years?

It is also important that some attempt should be made to anticipate what will next happen politically — and to guard against the over-optimism that present circumstances could engender. There is now talk of how the United Nations, with its solidity in the Middle East as a feather in its cap, might next move on to coerce countries that will not abide by international agreements to toe the line. This is an enticing concept. Would it not be marvellous if India, Israel and Pakistan could be forced to comply with the terms of the Non-Proliferation Treaty (NPT), and tomorrow? Of course it would. And would it not suffice that the permanent members of the Security Council should stitch together with the rotating members resolutions that would turn wishes into reality? That is a different proposition. Resolutions (as over Iraq) count for nothing until they are heeded. It might be different if the United Nations were more widely respected as an effective organization. That is another reason why this is a good time to put the technical agencies' affairs in good order. □

Greenhouse numbers

The use of a single number to assess the potential damage done by greenhouse gases is premature.

By what yardsticks will greenhouse warming be regulated if and when there is an international convention on the issue? The letter from David G. Victor on page 431 of this issue should serve as a warning to those eager to believe that the choice of regulatory yardsticks can be simple.

Although carbon dioxide is acknowledged as the chief potential cause of trouble, other gaseous materials such as the chlorofluorocarbons (CFCs) and methane will have similar effects on the Earth's surface temperature. So if there is to be a limit on national emissions of potentially dangerous materials, is it not best that the limits should be related directly to the potential effect of a nation's total emission of greenhouse gases? There would then be room for manoeuvre. Some might prefer to make a start by eliminating CFCs altogether, others might prefer to cut back on carbon dioxide right away. In any convention infringing on national sovereignty, flexibility of this kind, as in the decisions governments make about compliance with carbon dioxide quotas, can only help to make irksome agreements more acceptable.

Victor's argument bears on the concept of global warming potential (GWP), put into currency earlier this year by D.A. Lashof and D.R. Ahuja (*Nature* **344**, 529; 1990) and applauded in the draft report of the relevant working group of the Intergovernmental Panel on Climate Change. The idea is to represent the damaging potential of the various gases by a single number representing the integrated effect on surface temperature over the lifetime of the various gases in the atmosphere. CFCs and methane are notoriously more effective, molecule for molecule, than carbon dioxide, which is nevertheless produced so much more abundantly that its effects are likely to be predominant. In the calculation of GWP, the lifetime of the gas is crucial.

The essence of what Victor says is that the lifetime of carbon dioxide is conceptually still too uncertain for the relevant GWP to be calculated as a single number, let alone to be the standard by which the GWP of other gases is assessed. Lashof and Ahuja acknowledged as much in their original publication. Physically, the transfer of carbon dioxide either to plants or into the surface layers of the oceans can be relatively quick, but the processes governing transfer from the oceanic surface to the more commodious deeper layers are much more leisured, while carbon locked up in plants may be dumped back into the atmosphere in a century or less. As a consequence, it is not now feasible to strike a trade-off between a gram of CFC and a gram of carbon dioxide. Thus, while the sources of methane and the sinks for CFCs other than their mutual destruction with ozone are virtually unknown, the concept of a workable GWP should be a target for the future, not a yardstick for the here and now. □

Dutch cure for AIDS is discredited

- Faulty synthesis an "honest mistake"
- Eindhoven chemists at odds

The Hague

THE Netherlands research community has been set on its ears by an acrimonious scandal over the intended retraction of a paper from the Eindhoven Technical University, the exclusion of its principal author, Professor Henk M. Buck, from the chemistry faculty of which he was previously the dean and demands by the elected University Council for the resignation of the rector of the university, Professor M. Tels.

The disputed research report, published in April in the US journal *Science*, outlined a possible means of interrupting the replication of HIV in cases of human AIDS (*Science* 248, 208; 13 April 1990). The scheme, described as having been effective in *in vitro* trials, rested on the synthesis of oligonucleotides complementary to DNA in naturally occurring virus, but differing from natural DNA by the methylation of the phosphate group in each nucleotide.

It is now acknowledged by Buck and his colleagues that the synthetic route developed at Eindhoven for the construction of the methylated oligomers does not yield the products expected from it, and that the materials used in the *in vitro* trials contained little or none of the 20-nucleotide methylated oligomer supposedly designed to be complementary to crucial sequences of HIV.

While the organic chemistry in the collaborative research programme was carried out at Eindhoven, the biological assay of the synthesized material was the responsibility of Professor Jaap Goudsmit and two colleagues, human retrovirologists at AMC, Amsterdam.

Controversy began with the publication of the controversial paper last April, and appears to have been stimulated by a television broadcast on the same day in which Buck was shown walking his dog in the woods, and when claims were put forward that oligonucleotides methylated on the phosphate group would prove to be a means of halting AIDS infection.

Colleagues of Buck at Eindhoven, including one co-author of the article, Dr Leo H. Koole, then protested within the university that the claims were exaggerated and that the material used in the trials had not been fully characterized chemically.

A three-man investigating committee set up within the Faculty of Chemistry to investigate the allegations reached the conclusion, using HPLC and NMR mea-

surements, that they were correct. The committee's report to the rector also stated that Buck had been warned by his junior colleagues at Eindhoven that the materials used in the research had not been fully characterized, but that expressions of doubt had been brushed aside.

The internal report also said that Buck had been dictatorial in his management of the chemistry faculty, and that research other than that disputed might have been flawed by over-definite interpretation. On 30 August, the university said in a public statement that Buck's appointment as dean had been terminated. The decision that he should also cease to be head of the organic chemistry department was decided by a vote of his colleagues.

Speaking from his home earlier this week, Buck (who is 61) said that he considered the outcome of the investigation to have been unfair. While not disputing the technical conclusions of the three-man committee about the material synthesized at Eindhoven, he insisted that until then he had had no reason to suspect that the synthetic method would yield products other than those intended. "I was convinced about everything."

Buck also says that the method has been shown to yield the expected results when there are fewer than 10 nucleotides in an oligomer. What seems to have gone wrong is that a protective group used to shield reactive nucleotide sites behaves aberrantly when larger oligomers are synthesized.

On the charge that he had overridden scepticism of his work expressed by colleagues, Buck says that no colleague ever told him openly that the synthesis used was flawed. He says that only 25 per cent of his research time was taken up with phosphate methylation, and that his more than 300 published research articles showed that he had an important contribution still to make to organic chemistry. While acknowledging that the disputed research was in error, he believes he has been pilloried for making an honest mistake.

By coincidence, Buck and his immediate colleagues had made a careful presentation of the now-disputed research to this reporter, on a visit to Eindhoven early last November. At that stage, one member of the group was asked to display four test-tubes, each containing about 10 grams of a white substance, which were said to contain phosphate-methylated DNA oligomer tailored to the viral sequences of four identified AIDS

SOVIET ACCIDENT

Under a cloud

London & Washington

NEWS emerged last week of an explosion and fire at a metallurgical factory near Ust-Kamenogorsk in Soviet Kazakhstan, near the Chinese border, on 12 September, which is believed to have released a cloud of toxic beryllium oxide into the atmosphere. Local environment protection officers say that 120,000 people may have been exposed. Early reports suggested that the explosion had taken place at a nuclear facility but the International Atomic Energy Authority says the plant is not involved with nuclear fuel processing. Rather, it appears the plant was manufacturing beryllium for use in the nuclear industry.

It is not clear whether the industry is civilian or military. Beryllium can be used around nuclear reactor fuel rods to reflect neutrons back into the reactor core, but it is toxic and very difficult to work with. Other materials that are safer and easier to use are preferred for reactor use in the West, but beryllium is still used as a neutron mirror in nuclear weapons. It is processed in the United States at the Department of Energy Rocky Plate plant.

The beryllium oxide cloud is likely to have fallen to the ground quickly and over a relatively small area. The material causes fibrosis of the lungs and may also be a carcinogen. The local government has asked that the area be declared an ecological disaster zone but no details of the impact of the accident have been released.

Peter Aldhous & Alun Anderson

patients from Amsterdam.

Buck then also announced that a paper describing the findings had been sent to *Science*, where it was received on 10 October last. He now says that in the six months before publication, the journal had raised questions about the virology, but not about the organic chemistry.

The effects of this affair on the chemistry faculty and even the university have plainly been traumatic. Several members of the chemistry staff have left for posts elsewhere. One of those remaining says the past few months have been "like living in an unreal world", and that only now does he feel that "my two legs are coming back to the ground".

Wider issues have been raised in the University Council, some members of which hold that the university rector and his colleagues must shoulder the blame for having allowed Buck allegedly to tyrannize his academic colleagues. A particular concern that seems especially to have exercised council members is that Buck deprived a former colleague of research space, thus forcing him elsewhere.

The rector, Professor Tels, was not in Eindhoven earlier this week.

John Maddox

Not the last word

London

THE British government's long-awaited white paper (policy document) on environmental policy, released last week under the title *This Common Inheritance**, is under fire from opposition parties and environmentalist pressure groups for proposing few new initiatives to address global warming and other environmental problems. But Environment Secretary Chris Patten asserts that "this is not our last word on the environment", suggesting that the government may produce a more radical package in the future.

The white paper draws together existing initiatives in an overall strategy to protect the environment and promises that a minister from each government department will be responsible for its implementation. On the central issue of meeting Britain's stated target of stabilizing the emission of carbon dioxide at 1990 levels by 2005, the white paper rejects the introduction of a 'carbon tax' on fossil fuels in the next few years. A carbon tax is thought to be unacceptable politically to a government poised to sell the electricity industry to private investors as a high-profit concern.

In the short term, energy efficiency will be given a higher profile, under a new ministerial committee chaired by Energy Secretary John Wakeham, and the activities of the Energy Efficiency Office (which has seen its budget decline to £15 million in 1989-90, from the peak in 1986-87 of £26 million) will be stepped up.

A proposal from the UK Chemical Industries Association to replace the current multiplicity of UK pollution inspectorates with a single environmental protection agency, accountable to parliament, has not been accepted by the government. The association argues that a more single-minded approach would give the United Kingdom a stronger voice within the European Communities in framing environmental regulations.

The government has chosen to avoid this upheaval. But Her Majesty's Inspectorate of Pollution, responsible for policing industrial pollution to land and air, will be given a new advisory committee and greater independence from the Department of the Environment, in the hope that this will combat the inspectorate's widely publicized poor morale.

Patten was significantly less bullish at the launch of the new white paper than at the publication of the government's Environment Protection Bill in December last year. That bill created the foundations for pollution control in the United

Kingdom "well into the next century", he said. Last week, his vision was limited to laying the foundations of environmental policy "for the next few years".

But a close reading of the 300-page white paper points to a number of policies that may be introduced in the longer term. The potential for 'market instruments'

CARBON DIOXIDE EMISSIONS

Taxation or regulation?

London

THE decision of the UK government to avoid carbon taxes for the time being, announced in its new white paper (see above), will do little to settle the international debate over whether taxation or regulation will provide the best method to reduce the emission of greenhouse gases. In the United States and Europe, opinions now seem to be diverging on the most efficient policy instruments, and there are signs that different countries may head in quite different directions.

Stewart Boyle, from the UK Association for the Conservation of Energy, argues that regulations and increased subsidies to finance the initial cost of energy-efficiency improvements will be needed in addition to a carbon tax in order to meet the British target of stabilizing UK carbon dioxide emissions at 1990 levels by 2005. He points to studies showing that, during the 1970s, successive petrol price rises each decreased fuel consumption by British motorists for only a short period.

In California, opposition to a proposal to cut the state's carbon dioxide emissions by 40 per cent before 2010 (see *Nature* 347, 323; 25 September 1990) has centred around the size of carbon tax required to meet the target. Opponents of the move calculate that petrol (gasoline) would need to be increased in price by up to 70 cents per gallon and electricity bills by 20 per cent by 2000.

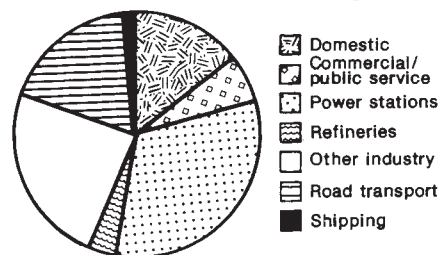
William Nordhaus, an economist from Yale University, says that both sides in the California debate are using "advocacy economics", rather than objective analysis. Looking at the United States as a whole, he says significant reductions will be costly whatever measures are taken, but Nordhaus maintains it is probably "twice as expensive" to achieve cuts in emissions through regulation rather than by taxation. The problem with regulations, Nordhaus explains, is that their effect is limited to individual sources of emissions, such as car exhausts, whereas carbon taxes can influence fuel use throughout the economy. Tax rises, in whatever form, are nevertheless so unpopular with the electorate that governments may prefer a limited package of regulations.

(such as the much vaunted carbon tax) in cutting pollutant emissions "will be pursued with vigour", and these are discussed in an annex to the main paper. Patten's 'special adviser', Professor David Pearce, an economist at University College London, has espoused these ideas for some time, but this is the first time they have received official recognition in a government policy document.

Peter Aldhous

Regulation to enforce energy efficiency standards is central to the Dutch strategy. The Netherlands is aiming to stabilize its carbon dioxide emissions at 1989-90 levels by 1995, with real reductions by the end of the century. Low-rate carbon taxes were introduced earlier this year (0.27 Dutch guilders on each 100 litres of petrol; G2.61 per tonne of coal), but the primary aim of these is to raise funds to finance environmental protection schemes and subsidies, rather than to have a real effect on fuel use.

In Britain, the merit of carbon taxes is not the only uncertainty surrounding the government's carbon dioxide emissions target — there is disagreement over the extent to which UK emissions will grow if no action is taken. The government's official position is based on a Department of Energy study which predicted an increase of 30 per cent by 2005. But this is widely dismissed as unrealistic, and the Department of the Environment is thought to be



Carbon dioxide emissions from the UK.

working on a 'business as usual' increase nearer to 20 per cent by 2005. Gerald Leach, who has completed an analysis of UK carbon dioxide emissions for the Stockholm Environment Institute, says that even this figure may be too high. Extrapolating forward from current trends in energy use and efficiency, he calculates an increase in UK emissions of less than 12 per cent between 1987 and 2005.

Peter Aldhous

Correction: Greenhouse gases

THE table in the 13 September news item "Room for improvement on emissions" gave total greenhouse gas emissions for a number of countries in millions of tonnes of carbon equivalent. The correct figures should be thousands of millions of tonnes. The per capita figures are correct.

*The white paper *This Common Inheritance* (Cm 1200) is published at £24.50 by HMSO, inclusive of a 36-page summary. The summary booklet is available separately at £2.50. Telephone orders accepted on 071-873-9090.

Sharing a shrinking budget

Washington

LAST week, the US National Academy of Science's Institute of Medicine (IOM) joined the debate over how best to support biomedical research with a 240-page call for a new balance in science funding*. Based on two years of meetings between some of the top names in biomedical research, the report takes the position that, given constant budgets, more money should be taken from research and put into training and facilities.

The IOM asked its panels to make their recommendations under some sobering assumptions. Three primary scenarios were stipulated: no real growth, 2 per cent real growth, and 4 per cent real growth in biomedical funding. Given these probably realistic constraints (congressional staff say biomedical research funding will be lucky even to keep up with inflation in the about-to-be-completed 1991 budget, despite an annual growth rate of some 10 per cent in numbers of new researchers), the panels chose to shift funding to students and laboratory construction and away from research. Floyd Bloom, chairman of the neuropharmacology department at Scripps Clinic, who led the IOM panel, says the members concluded that the National Institutes of Health (NIH) are "progressively under-funding the things that are necessary to keep the system vital". By funding established researchers at the expense of young investigators, "we're eating our seed corn", he says.

In the first two funding scenarios, the report calls for NIH to increase funding for training from 4.2 to 5.75 per cent of the grant budget by 1995 and to 6.75 per cent by 2000. Construction funds should also rise from 0.25 to 0.50 per cent of the budget over the next decade, the panels said.

Carol Sheeman, of the Association of American Universities (AAU), says her organization intends to "start beating the drum of this report" in lobbying Congress and NIH for reforms. "This is the clearest look I've seen at the issue of generational equity. It doesn't talk about redoing the whole system; it talks about adapting." Although Bloom says the panels expect a howl of outrage from established researchers, he points out that the proposals represent a total shift of only some \$20 million a year, or the loss of about 60 research grants out of nearly 5,000 in exchange for 400 new training awards.

Bloom also points out that much of the effect that the IOM panels were seeking can be achieved through private funding, rather than by cutting NIH grants. Indeed, the Howard Hughes Medical Institute (HHMI) is increasing its training budget by more than \$10 million this year,

to over \$65 million. "We are acutely aware that there is a problem with the pipeline, and we've made a decision to spend at least 20 percent of our funds on training and education", says Purnell Choppin, president of HHMI.

Another controversial aspect of the IOM report is the suggestion that grants be funded not necessarily at the full level requested, but at some level proportional to their relative ranking among other proposals. This "sliding scale" mechanism "takes into consideration the benefits or potential breakthroughs that could be derived from those grant applications falling below arbitrary cut-offs", the report says. For existing grants that face cancellation when they receive a ranking below the cut-off, the report suggests "step-down" funding, such as 60 per cent for an additional year.

Although many researchers rail at the idea of a cut in their grants, even if it means more for all, the proposals are likely to find a sympathetic ear in Congress. The report language in the current congressional appropriations bills emphasizes the importance of increasing the overall number of grants, at the expense of individual amounts, if need be. Responding to complaints of huge overhead, or "indirect costs", (up to 78 per cent) at many universities, Congress has also asked NIH to stretch the research dollar by favouring those with lower costs.

But NIH officials are concerned that such a move could hurt universities that are simply placed in an expensive part of the country. The consideration of indirect costs "would favour middle America schools over those on the East and West coast", says John Diggs, NIH extramural research director. AAU's Sheeman points out that schools in temperate climates would also be favoured because they spend less on heating and air-conditioning, two substantial portions of a institution's calculated overhead.

Nevertheless, NIH officials recognize that reforms, even drastic reforms, are inevitable as the number of researchers grows faster than the budget. "We have our work cut out for us", says Diggs. A series of NIH advisory panels are expected to finish their recommendations on the congressional proposals this month, and NIH is considering the IOM report as well. Given that every day without a 1991 budget (the fiscal year began this week with Congress still debating an agreement) means less money for all, NIH may be forced to impose some of the changes within months.

Christopher Anderson

*Funding Health Sciences Research — A Strategy To Restore Balance; National Academy Press, November, 1990.

Frontiers open up

Paris

THE Commission of the European Communities (CEC) has finally decided that the European Communities will take part in the Human Frontier Science Program (HFSP) for what is left of the programme's trial period (until March 1992). The decision follows a proposal put forward by CEC vice-president Filippo Maria Pandolfi and will allow participation of several European countries that have so far not had full access to the programme.

HFSP grew out of plans first put forward officially by the Japanese government at the 1987 Venice economic summit of the group of seven (G7) leading industrial nations. For a three-year trial period, Japan is putting up the lion's share (¥2,384 million for the first year) of funding for an international, interdisciplinary basic research programme into higher brain functions and biological functions at the molecular level. The programme has given scientists from the G7 founder-member countries access to a new and attractive source of funding.

Janet Watford, responsible for HFSP within the CEC, explains that Japan also invited the European Communities to take part, as they are also represented at G7 summits. But official procedures have meant that, until the recent CEC decision, European member states (apart from Britain, France, Italy and West Germany) have not been able to participate fully in HFSP. Now, says Watford, smaller members of the European Communities will be able to take part on the same basis as the founder members and can be principal researchers on joint projects.

CEC will now provide bursaries and organize workshops for HFSP research, but no new money will be available. Instead, existing funds ear-marked for CEC research programmes will be made available for applications under HFSP's terms of reference. P.C.

HIGH-ENERGY PHYSICS

A proton shared...

London

MUON research is to be the latest beneficiary of Japanese research spending in the United Kingdom. The Japanese Institute of Physical and Chemical Research (RIKEN) will spend £6 million to set up a complex of muon beamlines on Isis, the spallation neutron source at the Science and Engineering Research Council's Rutherford Appleton Laboratory. Researchers from Japan and Britain will use the new facility, studying muon-catalysed nuclear fusion and measuring magnetic fields at the atomic level. RIKEN will save money through the collaborative arrangement, dispensing with the need to build a source of high-energy protons (which Isis produces) to generate muons. P.A.

Seismic warnings hit home

Buffalo, New York

THINK of earthquakes in the United States, and you think of California. At least Californians do, and they still react angrily to the mention of a \$50 million earthquake research centre on the East Coast campus of the State University of New York at Buffalo (SUNY). Nearly five years after the National Science Foundation (NSF) snubbed California by making New York the National Center for Earthquake Engineering Research (NCEER), West Coast researchers hint darkly that the NSF's decision was due more to the fact that former NSF director Erich Bloch is a SUNY-Buffalo alumnus than to any great need for an East Coast earthquake laboratory.

But over the last few years, NCEER has turned out to be adept at doing one thing that might well have been beyond the abilities of any California laboratory: convincing the East Coast political establishment to take earthquakes seriously.

Due in large part to efforts by NCEER researchers, Connecticut and New Jersey now require new buildings to be earthquake resistant, as does the city of Memphis, Tennessee, and its surrounding county. The state of New York is considering new measures. And now New York City, after long consultation with NCEER, is preparing a new building code that would require buildings to be able to withstand an earthquake of up to magnitude 5.5 on the Richter scale. "Once New York City goes, the rest of the East Coast will follow", predicts civil engineer Carl Costantino, a NCEER collaborator at the City University of New York.

The chance of an earthquake like last year's San Francisco tremor (magnitude 7.1 on the Richter scale) hitting the East Coast is slim, by most estimates, but not as slim as it was once thought.

Recent research by the US Geological Survey suggests that there is a 40–60 per cent chance of an earthquake of at least magnitude 6.0 in the eastern United States before 2020 (see *Science*, 249; 21 September 1990). The most violent earthquake in US history took place in the New Madrid area of Missouri in 1811, and Charleston, South Carolina, was virtually levelled in 1886.

In the New York City area alone there is a better than even chance of a magnitude 5.0–5.5 earthquake in the next 10–15 years, according to NCEER researchers. Although an earthquake of that magnitude is unlikely to damage skyscrapers and other large modern buildings, smaller and older brick or stone structures could collapse without warning. One study suggested that a magnitude 6.0 earthquake in the New York City area could cause direct damage of \$11,000–\$25,000

million, neglecting fire and human costs.

A NCEER team that visited Soviet Armenia shortly after the catastrophic earthquake there in 1988 found that simple techniques such as reinforcing roof-wall connections and inserting steel rods into the walls could have prevented much of the destruction. In the United States, the researchers began to publicize parallels between eastern US cities and those in Armenia, and found a receptive audience. As NCEER executive committee member Klaus Jacob put it, "a few pointed prime-time evening news TV essays finally broke the ice". New York's building commissioner promptly called for an expert committee to draft seismic provisions.

Since then NCEER researchers have campaigned tirelessly, using every opportunity to drive the point home. They assembled a team to visit San Francisco

NATURAL HISTORY MUSEUM

Who will buy?

London

LONDON'S Natural History Museum, apparently undismayed by criticism of its corporate plan, is distributing its latest mail-order catalogue in good time for the Christmas trade. Many Londoners found the 32-page colour catalogue falling from their newspapers at the weekend.

Called the *Catalogue of the Unusual*, the document features over 160 gift ideas from the inexpensive (a set of car stickers at £2.95) to the extortionate (limited-edition dolphin sculptures at £145); the frankly tasteless (both the above) to the arguably exquisite (illustrated playing cards at £14.50 per set); the frivolous (armadillo-shaped sleeping bag, £99.50) to the ingenious (an electrolytic digital clock powered by two potatoes, £12.95); and the indulgent (silk tie with M. C. Escher design) to the practical (the prizewinning 'green lensman' portable field microscope, with optional SLR adaptor and case for a very reasonable £122.90). There seems to be something for everyone.

The traditional festive plastic dinosaurs (now relegated to a subfusc item on page 26) are largely replaced by a superabundance of designer-print leisurewear and silver jewellery. This marks the transfer of the museum's mail-order business to an external company, Design Marketing Ltd of Andover, Hampshire.

The concept follows similar ventures by other museums such as the Smithsonian Institution in Washington, DC, (whose sculpture of a group of meerkats was a bestseller, and is now available from the Natural History Museum catalogue at £32.95). The American Museum (Natural History) has a nice line in dinosaur-print silk ties.

Henry Gee

after the 1989 earthquake and sponsored the travel of four members of New York City's code-writing committee to accompany them. The committee recommended a seismic code almost immediately on their return.

NCEER officials are currently considering a small demonstration project in New York City for the centre's advanced technology, such as sliding foundation bearings and active controls that use movable weights or hydraulic braces to counter the forces generated by an earthquake or strong winds. And next year, in the first such US project, NCEER researchers will equip an existing skyscraper in the American southwest (details of which have not yet been released) with wind-resisting active controls. If successful, the demonstration will no doubt go far in convincing the construction industry to consider similar measures in other hurricane- and earthquake-prone areas.

Christopher Anderson

USSR—US UNIVERSITY

Academic exchange

Moscow

THE first 60 students have now been enrolled in the independent Soviet-American University, which held an opening ceremony here last week. The objective is to exchange teachers, rather than students, between the Soviet Union and the United States, perhaps economizing on costs but also providing a novel environment for teachers.

The university, which has the backing of President Mikhail Gorbachev and President George Bush, is the concept of people such as Yuri Osipyan, vice-president of the Soviet Academy of Sciences, and Professor Edward Lozanski from Washington, who said at last week's ceremony that the project "seemed like science fiction 12 months ago".

For the immediate future, the two sides will exchange teachers in the fields in which they are themselves strong. The Soviet side will at first offer experts in, for example, theoretical physics, analytic mathematics, Slavonic studies and music education. The United States will send experts in computer science, economics, business and management.

The new university will be housed initially at the Moscow State University, which will be responsible for awarding qualifications. The plan is that the university, when independent, will offer four-year courses, as in the United States. Funds are being provided by the USSR Academy of Sciences, Moscow State University and the USSR State Committee for Public Education, as well as by several private funds, foundations, companies and independent organizations in the United States.

Valeria Prut/
Novosti, Moscow

Europe in sad decline

London

EUROPEAN semiconductor researchers say that last month's decision by the Dutch electronics company Philips to disband a research group studying gallium arsenide heterostructures, is part of a continuing wave of cutbacks in support for long-term research on semiconductors by European companies. Coming as it does after recent cuts in semiconductor research by the British companies GEC and Plessey, this decision means that semiconductor groups at UK universities must now consider forging closer links with Japanese companies which can still afford to expand their research efforts.

Almost half of the 390 staff at Philips Research Laboratories in Redhill, Surrey, will be shed during the next two years, including the 15-strong group working on gallium arsenide semiconductors. Bruce Joyce, director of the Science and Engineering Research Council (SERC)'s semiconductor materials Interdisciplinary Research Centre at Imperial College, London, says that the loss of the gallium arsenide group is a "serious blow". The Philips researchers collaborate extensively with research groups in British universities.

Both Joyce and Professor Colin Humphreys, from the University of Cambridge, who is chairman of SERC's Materials Commission, say that major European companies seem to regard gallium arsenide as a perpetual 'tomorrow's material', which may never justify a large research and development effort. At the same time, Japanese companies are invest-

ing heavily, to exploit the potential of gallium arsenide to give faster electronic devices and to develop its electrical response to light as the basis for an optical electronics industry.

SERC currently spends about £3 million a year on the semiconductor centre at Imperial College and its low-dimensional structures and devices initiative, in addition to strong research grant support for semiconductor research. This falls within SERC's category of strategic research, where future commercial applications are expected. Traditionally, this spending has been justified as underpinning British industry. But Humphreys says this has become "a very dangerous argument" for semiconductor research, with the British and now the European industrial research effort declining.

SERC's expenditure on semiconductor research may in future aim at providing a strong British research base to attract overseas investment, rather than at supporting British industry. There are signs that this is already happening. The Interdisciplinary Research Centre at Imperial College agreed a five-year collaborative research project on semiconductor films with the Research Development Corporation of Japan earlier this year, which will bring in several million pounds of Japanese money (see *Nature* 343, 300; 1990); and the Japanese electronics company Sharp is to open a new research laboratory in Oxford in 1992, eventually employing 100 scientists including an optical electronics research team.

Peter Aldhous

Scientists of note

PAUL Ehrlich (1854–1915), the German immunologist and founder of chemotherapy, is to be immortalized on this new 200-mark banknote, which will be issued on 1 October by the Deutsche Bundesbank. Alongside his portrait, at left, is a stylized rendition of arsenobenzene, a precursor for the syphilis drug salvarsan which Ehrlich developed. Ehrlich worked in Berlin and Frankfurt and won the Nobel Prize in 1908.

Ehrlich is not the first scientist to be honoured on a German banknote; nineteenth century chemist Justus von Liebig appeared on a one-hundred Reichsmark note issued by the Third Reich in 1935.

Other scientific figures to have appeared on notes in this century include Louis Pasteur and Blaise Pascal in France, Isaac Newton in the United Kingdom and Alessandro Volta, who appears on the 10,000-lira note now in circulation in Italy.

S.D.



New bid for fusion

Tokyo

JAPAN'S Science and Technology Agency (STA) has decided to join the crowd by putting in a bid to base the thousand-million dollar second phase of development of the International Thermonuclear Experimental Reactor (ITER) in Japan. Nine US regional consortia, the European Communities, and the Soviet Union are also chasing the five-year engineering design phase of the project due to begin next year (see *Nature* 347, 114; 13 September 1990). But although Japan is not alone in the international race for ITER, it has come up with a unique domestic reason for seeking funds for the project.

The STA has asked for more than \$60 million for ITER and other activities from an unusual budgetary account, available for one year only and intended to "improve the lifestyle of the Japanese people". Most of the general accounts for the various ministries are being held steady and some, like that for the Ministry of International Trade and Industry, are even contracting. So government officials welcome the special 'lifestyle' fund set up after the United States put pressure on Japan to ease trade tensions by boosting domestic spending.

STA hopes that lifestyle funds will both support ITER and help to repair old facilities at national research institutes. STA officials argue that better national research institutes will improve the lifestyle of the Japanese people through the development of new technology, while fusion — that is, ITER — is the "ultimate energy source" which will benefit all mankind.

But Ikuo Makino, of STA's Atomic Energy Bureau, admits that once the Ministry of Finance makes a decision on the lifestyle budget requests at the end of December, any extra money received will simply be "absorbed" into STA's ordinary budget and "everyone will stop talking about improving Japanese lifestyle". D. S. US TAXES

Small benefits

Washington

AMONG the tax increases and levies in the new US budget agreement, which was settled in a late compromise this week, is a rare bit of good news for science. Small corporations (those with total assets of less than \$50 million) will be allowed to increase the size of their research and experimentation tax credit from 20 to 30 per cent. They will also be able to take a tax credit on scientific equipment, by writing off its total depreciation in the first year, effectively lowering the cost of the equipment and encouraging expansion. Budget negotiators estimated that the two provisions would be worth \$400 million to US companies. The agreement must yet be ratified by Congress before it becomes law.

C.A.

Levelling the field

Paris

FIFTEEN US delegates from different government departments last week met members of the council of the European Space Agency (ESA) in Paris to discuss the ground rules for commercial satellite launches. The aim is to put an end to a long dispute about the legitimate role of state development funds for commercial rocket launchers.

For the past six years, US commercial launcher manufacturers have complained that Europe's Arianespace is competing unfairly, using state money (through ESA) to develop its rocket launchers. In response, Arianespace has pointed out that the US shuttle was developed by NASA and that



defence contracts keep the commercial launcher manufacturers in business.

Last week's meeting was the first of a long round of talks intended to produce a united front against launches offered by China and the Soviet Union. A launch aboard the Chinese Long March rocket can cost half as much as an Ariane launch.

An ESA spokesman said after the meetings that further discussions would be needed before any decisions on the definition of fair play would be made. P. C.

Fire down below

Tokyo

JAPAN'S National Space Development Agency (NASDA) is having no luck at all with development of the main engine of its next generation H-II rocket. Last week, the engine again burst into flames during a test run at NASDA's Tanegashima Space Center.

The fire, the latest of a string of mishaps that have plagued the engine since tests began last year, appears to be the result of new problems. It broke out around the oxygen turbopump 16 seconds into a 200-second test firing and, although it was extinguished in about 20 minutes, it caused extensive damage to the engine.

NASDA officials are not yet prepared to say what effect the latest fire will have on the H-II launch schedule. But its effect is "under investigation" and Ryuichi Nagashima of NASDA's Program Planning and Management Department says the fire is "serious" because "we are working on a very tight schedule". D. S.

Study will break new ground

London & Washington

UNITED States health officials gave their seal of approval last week to the most ambitious long-term study of the AIDS epidemic in North America ever conducted. As well as assessing the effectiveness of approved therapies such as AZT, the study will monitor the use of 'underground' drugs and alternative therapies such as acupuncture and Chinese herbs. Hundreds of physicians and thousands of patients at over 40 community-based clinics in the United States and Canada are likely to participate.

In announcing the project, the Department of Health and Human Services Secretary, Louis Sullivan, whose past policies have been bitterly criticized by AIDS activists, said it "will result in the most comprehensive collection of information on HIV-infected persons amassed in the history of the AIDS epidemic". The study will be sponsored both by the National Institute of Allergy and Infectious Disease (NIAID), which coordinates the bulk of state-funded AIDS research in the United States, and the American Foundation for AIDS Research, a private funding organization.

The central aim is to gather data showing how AIDS progresses in different patient groups. So far, attention has focused largely on gay and bisexual men with AIDS, but the new study will look at a broader spectrum of the patient population including women, pregnant women and intravenous drug abusers. The study's designers hope that ultimately it will provide a detailed picture of the kinds of opportunistic infections affecting different patient groups and whether the pattern of infection is shifting with time.

Doubts over 'wonderdrug'

London & Washington

THE African AIDS drug Kemron has touched off a heated debate with a racial twist. Hailed as a 'wonderdrug' by its proponents, but considered worthless by sceptics, Kemron was given an optimistic launch earlier this year by Kenya's president Daniel arap Moi. Now it is in demand in the United States.

Kenyan scientists say their trials of the drug, whose activity they attribute to a tiny dose of α -interferon, reveal marked benefits in over 1,000 patients, including 50 who they say are no longer HIV-positive. But tests of low doses of α -interferon in the United States have yielded a mixed response. Some patients have reported no benefits; others claim modest to substantial improvements. Past US research on the effects of high doses of α -interferon has proved equally inconclusive.

There is, however, another dimension

They also hope to analyse mysterious regional variations in opportunistic infections.

Anthony Fauci, director of NIAID, is boldly optimistic. The study should "help to identify areas where new research is needed, and may result in some new approaches to treatment", he says. But Martin Delaney of the AIDS organization Project Inform, which will participate in the study, thinks expectations are "higher than realistic". It will provide useful information on drug interactions and safety, he says, but would have had more impact five years ago.

The use of approved experimental drugs currently in clinical trial, such as ddI, ddC, interleukin-2, α - and β -interferon and soluble CD4, will be monitored. Some of these are available through underground networks and ddI and ddC are likely to be distributed through official expanded access drug schemes.

Another area of interest is the use of alternative therapies and unorthodox drugs. Even though the study will provide only a rough guide to their effectiveness, anything "distinctly detrimental or miraculously favourable to the patient will show up", says Mathilde Krim, director of the American Foundation for AIDS Research.

Information on unorthodox drugs will be keenly awaited. With such drugs becoming increasingly available from underground markets, drug fads influenced by anecdotal reports of 'cures' abound. Compound Q — a drug claimed by some to have remarkable therapeutic effects — hit the headlines earlier this year when two patients died using it.

David Concar & Diane Gershon

to the debate. Black health officials in Africa and the United States say the drug has been played down by the white-dominated medical establishments of the West because of an ingrained distrust of African science.

This allegation has been flatly refuted by US health officials who say they are keeping an open mind on the matter. The chief problem is that the use of two different forms of the drug — powder and tablet — and a lack of placebo controls have rendered the data from African trials almost impossible to interpret.

Sadly, the confusion shows no sign of going away. Last month, a meeting convened by the World Health Organization specifically to discuss Kemron concluded that it is still too early to tell whether low doses of α -interferon are of any value to AIDS patients.

David Concar & Diane Gershon

NUCLEAR WEAPONS

Democratic Brazil says no . . .

São Paulo

BRAZIL's president, Fernando Collor de Mello, last month destroyed a secret nuclear test facility that his military predecessors had built in the Amazon and so quashed all rumours that Brazil intends to develop nuclear weapons. Before setting off on a trip to Europe and the United States, where he was to address the United Nations assembly, Collor went to the site where a 320-metre-deep shaft had been drilled in preparation for an underground nuclear test and sealed the shaft for ever.

In keeping with his media-conscious style, Collor dropped two shovels full of lime into the pit, located in the Cachimbo mountain range, in the state of Pará, near the border with the states of Mato Grosso and Amazonas, deep in the rain forest. The secluded nuclear weapons test site was built in the early 1980s, and remained secret until newspaper reports began to appear in 1986. The Cachimbo proving ground is managed by the Brazilian Air Force.

The Air Force, along with the Navy and Army, had nuclear research projects independent from the government's official programme to develop nuclear power reactors with West German help.

NON-PROLIFERATION TREATY

But the German agreement came to nothing, and the military met with mixed results. The Navy is far ahead of its sister services, with a uranium enriching plant in Iperó, in the state of São Paulo. The Army programme is, however, more worrying, because while the Navy says it needs plutonium to power a nuclear submarine, the army can want plutonium only to make a bomb.

Physicist José Goldemberg, secretary for science and technology and long-time opponent of the nuclear deal with West Germany, said that Collor's action showed once and for all that Brazil does not want to explode nuclear devices. But the controversy is not yet over. It is now becoming clear that someone in the recent past had intended to build nuclear devices, given that they had gone as far as to prepare a place for the tests. The opposition parties are now demanding to hear the whole story.

The best known 'green' representative, São Paulo's social democrat Fabio Feldmann, says that he intends to find out what happened even if he has to sue the armed forces on the ground of "contempt of Congress", an action which the new constitution makes possible.

Ricardo Bonalume

. . . South Africa thinks again

Cape Town

SOUTH Africa's Minister of Foreign Affairs Pik Botha said last week that South Africa may at long last be prepared to sign the Nuclear Non-Proliferation Treaty (NPT), provided that other states in southern Africa will also do so. Botha's statement, made on the first day of the annual meeting of the International Atomic Energy Agency (IAEA) in Vienna, was the most explicit commitment yet by the South African government to sign the NPT and succeeded in staving off a resolution to expel South Africa from the agency, for the fourth consecutive year.

The statement will help the African states who pressed for the continent to become a nuclear weapon-free zone at the recently concluded NPT conference in Geneva.

Mozambique began the process of accession to the NPT on 12 September, and diplomatic efforts are now being made to coordinate a commitment to sign by Zimbabwe, Zambia, Angola, Namibia and Tanzania. Botswana, Lesotho and Swaziland are not considered to have the potential to develop nuclear weapons.

Botha also said that talks on a comprehensive safeguards agreement for South

Africa's nuclear facilities may soon be concluded. South Africa's single nuclear power station at Koeberg, outside Cape Town, has been under IAEA inspection since it began operating in 1984, but the new commercial uranium enrichment plant at Valindaba, which supplies the reactor's fuel, is not.

A separate pilot enrichment plant of Valindaba has caused most international concern because it has been thought capable of producing the highly-enriched uranium needed for nuclear weapons. But the plant was shut down last February and has now been dismantled.

The chief executive of the Atomic Energy Corporation, Waldo Stumpf, denies that the closure of the pilot plant was linked to the NPT negotiations, and says that the plant was no longer needed because sufficient material for the country's SAFARI-I research reactor (also under IAEA inspection) has been produced, and because the commercial plant is now running satisfactorily. The material from the pilot plant is stored could be inspected by IAEA, but the possibility remains that it does not contain all the nuclear material produced during its lifetime.

Michael Cherry

ENDANGERED SPECIES

Pyrenean bears bait French hunters

Paris

JUST weeks after France ratified the 1979 Bern Convention on the protection of endangered species of plant and animal, the authorities already have a fight on their hands over the fate of the Pyrenean bear — one of the 680 endangered species listed in the convention.

Bear-hunting has been outlawed in the French mountains since 1962. But local farmers — who consider the animal a pest — have been able to hunt its principal prey, the wild boar, and to fell trees, destroying its habitat. As a result, the population of Pyrenean bears has fallen from 170 in 1940 to 10–20 today. Under the terms of the convention, protection must be given to the habitat of listed species. So, despite fierce opposition, the secretary of state for the environment, Brice Lalonde, has banned all hunting from a 6,500 hectare zone where up to 13 bears, including one cub, live.

Lalonde has also offered the forestry

Ediciones Sicilia

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Under threat — protection measures for the Pyrenean bear may not be enough.

office compensation to delay logging in this area for a year. The move has outraged local farmers. Some would be happy to see the bears disappear altogether. Every year, they say, sheep, goats and cattle are destroyed by bears. But, as the hunting season opens, the nation's two million bloodsports enthusiasts — who will go for everything from thrushes to deer — are also angry because they believe that their chances of bagging wild boar will be hampered by the new sanctuary zones. Environmentalists want to make sure Lalonde enforces the new measures but, given the opposition, it seems unlikely that he will go the whole hog and introduce new bears to the region, as the Bern convention proposes.

Peter Coles

MRC and peer review

SIR—As a clinician, I am appalled by the Medical Research Council (MRC)'s lack of support for work that has direct potential for advances in the clinical care of patients.

The question that Dr David Rees does not answer in his response to your leading article "Untransparent grants" (*Nature* 346, 684; 1990) is "How does the MRC measure the merits and demerits of individual proposals?" The traditional mechanism is peer review: literally, judgement by equals, which in this context must mean authorities with an equal ability to judge the science under review — experts in the specifics of the subject. Although one would not normally expect this to be so, in the actual case that sparked off your leading article, there was unanimity of view among the peers — that the science was of high quality and of clinical importance, and should be supported — yet the MRC rejected this judgement.

Clearly, the MRC is unable to finance, at the level that the science itself justifies, all the work that is highly rated by peers. Rather than spread the hardship evenly, the MRC has chosen to deal with this situation by discontinuing excellent teams in some fields while sparing others in a harsh all-or-none policy. If the MRC is now saying that such decisions *between fields* are based on measurements of the quality of science in each field, then we need to know how these extraordinarily difficult measurements are made and what precision they are thought to have. Many will doubt that such measurements are possible: these are subjective value judgements, not precise measurements of physical properties. Moreover, logic dictates that what is excellent in one field remains excellent, whatever the competition for funds.

Clinical trials

SIR—I would like to correct statements in the News story headlined "Clinical trials" (*Nature* 345, 754; 1990). The policy recommended by the Advisory Committee on Women's Health Issues (ACWHI) on including women as study subjects was published in the National Institute of Health (NIH) Guide to Grants and Contracts in 1986 and again in 1987. The article mentions two internal surveys indicating that sexual imbalances existed in research within NIH, following which ACWHI sent new recommendations to the director of NIH. I am unaware of any internal surveys on this subject during this period.

The ACWHI did submit a report on NIH support for research on women's health issues (defined as diseases and conditions unique, more prevalent, more

I believe that the judgements now being made are not of this sort at all, nor are they peer judgements, but management decisions taken to achieve specific policy objectives within a given financial structure. This has now been admitted by Rees himself, in his interview with the *British Medical Journal* (301, 508; 1990) when he makes it clear that it is now MRC policy to throw the peers' opinions out of the window if the nine-man MRC strategy committee so decides.

The MRC is a publicly accountable body, and the manner in which its funding decisions are now being made should be approved by the public. If the present situation continues, not only the reputation of the MRC, but also that of British medical science, will certainly suffer.

ROY CALNE

University of Cambridge Clinical School,
Department of Surgery,
Level 9, Addenbrooke's Hospital,
Hills Road, Cambridge CB2 2QQ, UK

SIR—Please add my voice to those protesting at the forthcoming closure of the MRC Medical Cryobiology Unit in Cambridge (*Nature* 364, 685; 1990).

Although I know little of the circumstances influencing the decision to close this unit, the council's other research priorities or the marketability of the work undertaken in Cambridge, I am familiar with both the Cryobiology Unit in Cambridge — having undertaken a 'sandwich' year there as part of my degree course — and cryobiology internationally. I can therefore appreciate the high esteem in which the work performed at Cambridge is held by the 'cryobiology community' and the long-term (but perhaps not immediate) benefits of the pioneering research into several aspects of cryopreservation

serious, having different causes or interventions in women or some subgroup of women) in August 1989, the month after Dr Wyngaarden left. The committee recommended that the policy published for grants and contracts also be applied to the NIH intramural research programme. Unfortunately, that recommendation was not brought to the attention of the responsible officials. Although data show equal access for men and women to research conducted in the clinical centre at NIH, adoption of a stated policy on including women is now under discussion.

IRIS J. SCHNEIDER

(Assistant Director for Program
Operations and Planning, NCI &
Co-Chair, NIH Advisory Committee
on Women's Health Issues)

National Institutes of Health,
National Cancer Institute,
Bethesda, Maryland 20892, USA

performed there.

My own time at the unit was most constructive, gave me a sound base in science and an invaluable launch to my career.

M. J. BENTON

Rue de la Poste,
St Ouen les Vignes,
37530 Amboise, France

No cancer link

SIR—In a recent News article¹, reference is made to a draft report from the US Environmental Protection Agency (EPA) concerning a supposed link between exposure to extremely low-frequency (ELF) electromagnetic fields (EMF) and cancer. The article quotes from the report: "The consistently repeated pattern of lymphoma, leukemia and brain cancer in the childhood studies and the ruling out of several confounding exposure factors in the Savitz *et al.* study (*Am. J. Epidemiol.* 128, 21; 1988) argue in favour of a causal link between these tumor types in children and exposure to ELF magnetic fields or electric fields." There are several misconceptions in this statement.

The inconsistency of findings reported by investigators working on this topic does not support the argument of a causal link. Regarding possible confounding exposure factors, Savitz and Feingold² found that the incidence of childhood cancer was associated with traffic density. Increased risks for total number of cancers and leukaemias were related to increased traffic densities. The data were obtained during the *same* study of EMF and cancer cited by EPA. The odds ratios for these associations were *greater* than those reported earlier for EMF and cancer. In another study often quoted as being evidence of a link³, "cases were found to generally live closer to high traffic routes". One potential consequence of high traffic density is a high level of benzene⁴, which is a well-established cause of leukaemia.

Cartwright⁵ summed it up best when he stated: "Our present scientific knowledge points at the very best to a minute risk of EMF verging on the point of non-existence." The unsubstantiated claims of a link between EMF and cancer should not lead readers to believe that a hazard exists.

JAMES R. JAUCHEM

Radiation Sciences Division,
US Air Force School of Aerospace
Medicine,
Brooks Air Force Base,
Texas 78235-5301, USA

1. Shulman, S. *Nature* 345, 463 (1990).
2. Savitz, D. A. & Feingold, L. *Scand. J. Work Environ. Hlth* 15, 360–363 (1989).
3. Wertheimer, N. & Leeper, E. *Am. J. Epidemiol.* 109, 273–284 (1979).
4. Stenberg, U., Alsberg, T. & Westerholm, R. *Environ. Hlth Perspect.* 47, 53–63 (1983).
5. Cartwright, R. A. *Br. J. Cancer* 60, 649–651 (1989).

Vote of no confidence

SIR—The crisis at London's Natural History Museum, which erupted when the corporate plan was announced on 23 April this year, has now lasted more than four months and shows no sign of ending. The essence of this crisis is that the plan will result, whatever its stated intentions, in narrowing the span of taxonomic and systematic research in this museum. And, yet, as one of our trustees, R. M. May, writes (*Nature* **347**, 130; 1990), "without taxonomy to give shape to the bricks, the systematics to tell us how to put them together, the house of biological science is a meaningless jumble".

About a thousand letters of protest have been sent to the relevant minister by our colleagues from all over the world who recognize that this museum is the world centre for taxonomic expertise (*Hansard* 177 (148), col. 18; 1990); Mr Tam Dalyell MP has several times objected to the plan in parliament (*Hansard* 174 (118), cols. 77–81, 1990; *Hansard* 174 (126), cols. 1084–1092, 1990; *Hansard* 175 (135), cols. 1025, 1075–1077, 1990); there have been two days of strikes; and there has been a storm of press comment, nearly all of it critical of the plan and including a swinging leading article in *Nature* (**346**, 397; 1990). But the net result of all this protest, so far, has been that three scientists threatened with forced retirement have been temporarily reprieved — two in palaeontology and one in zoology. This is from a total of some 50 jobs under threat.

There have been no reprieves in mineralogy, botany, entomology or the library nor any significant dialogue or negotiation between management and staff. A new management structure, with imposed separation of curation from research for some 150 people, has been forced down our throats, as has also a brutal system of short-term contracts for researchers. And our prizewinning design team is still threatened with extinction. Moreover, the director's main response to the letters of protest is blandly to point out their usefulness in the search for funding, since they demonstrate that the taxonomic community of the whole world is interested in the fate of the Natural History Museum (*Museums Journal* **90** (8), 8; August 1990). He is right about that. But the barrage of letters also shows that the world taxonomic community has no confidence in the director nor in his plan. The scientists and designers employed in this museum similarly have no confidence in him.

We implore the new Minister for Arts, Mr David Mellor, to reject the draft of the corporate plan as now submitted to him. As motivated and competent scientists who have been working here for many years, we have a right to be heard on

questions of the organization of science in this museum. Our scientific attainments and qualifications are more impressive than those of the top management.

COLIN PATTERSON (Chairman), R. P. S. JEFFERIES (Secretary), K. SATTLER, P. WHEATCROFT, (Branch Chair), JULIET CLUTTON-BROCK, C. J. HUMPHRIES, C. R. HILL, P. H. GREENWOOD

*Science Defence Committee,
Natural History Museum,
Cromwell Road,
London SW7 5RD, UK*

SIR—Beverly Halstead of the Geologists' Association (*Nature* **346**, 406; 1990) accepts a number of points I made about the Natural History Museum. I wish to make it clear that I am a passionate supporter of the importance of museums in society and that I am disturbed at many of the events that have occurred recently in Britain.

But it is vital to understand that many museums now acknowledge that their public educational role must be carried out in a different manner from what it was some decades ago, and for good reason. Moreover, there can be no special reason why museum people should not change the way they work when change, some of it horrendous, is everywhere about. Indeed, there have been severe shortcomings in the effectiveness of the way in which many museum people have worked in the past — trustees and directors as well as security and cleaning staff.

John Evans in his letter (*Nature* **346**, 406; 1990) misses my point about Disneyland, which is described in glowing terms as a place to work and as a place to visit. Could one really have said that about museums of the past as frequently as one would have liked?

Finally, I note that 100 staff posts are to be lost of which 51 are scientific: is anyone worried about the other 49 posts or people even perhaps? I haven't seen a whisper about them.

DES GRIFFIN
(Chairman, Council of Australian
Museum Directors)

*c/- The Australian Museum,
6–8 College St,
Sydney, NSW 2000, Australia*

SIR—I was shocked to hear about the decision to reduce the staff at the British Natural History Museum and the accompanying reshaping of its research programme. That this should happen in the face of the fact that scientists all over the world look to the British Natural History Museum as the repository of the finest collections of objects of natural history and to its staff for expert help in the solution of their research problems makes the

contemplated change even more reprehensible.

In my own branch of study, palaeobotany, the museum has one of the finest collections of fossils in the world. Thinking that the museum was the best place for them, I handed over many of my type slides and specimens and also the figured duplicates to the museum.

In this connection, I must point out that the museum's collections belong not only to Britain but also to the entire world, irrespective of whether the material is of British origin. The museum's founders never imagined that the control of the museum would one day pass into the hands of short-sighted people who would regard its research as esoteric and peripheral. Small minds sometimes take control of things that are too big for their comprehension and, power-drunk, they start destroying, right and left, our gifts of the ages. However, the innate vitality within our heritage and also that of the general public should be able to assert itself and check such arbitrary and thoughtless ruination.

DIVYA DARSHAN PANT

*Botany & Palaeobotany Laboratory,
University of Allahabad,
106, Tagore Town,
Allahabad 211002, India*

Cell line ownership

SIR—Your report on the California Supreme Court's decision of 9 July 1990 rejecting a patient's lawsuit to recover profits from research on a cell line derived from his surgically removed spleen tissue ("Tissues not for sale," *Nature* **346**, 208; 1990) contains several inaccuracies.

First, although the University of California has a patent on the cell line, no licence has ever been granted on the patent, to Genetics Institute, Sandoz or any other party. Second, I received stock in Genetics Institute as compensation for my work over a seven-year period as a founding scientific consultant to the company, not in return for providing access to the cell line. The university, which owns the cell line, arranged with Genetics Institute to provide access to the cell line before a patent was issued, in exchange for research funding in a collaborative research agreement. Since the issue of the patent in 1984, the cell line has been available to any qualified investigator from the American Type Culture Collection, Rockville, Maryland. However, because the cell line produces the human retrovirus, HTLV-II, it must be treated as potentially hazardous material.

DAVID W. GOLDE

*UCLA School of Medicine,
Center for the Health Sciences,
Los Angeles, California 90024, USA*

The Gulf crisis and research

SIR—The world community has been almost unanimous in condemning the Iraqi invasion and occupation of Kuwait and in demanding the immediate withdrawal of Iraqi forces from Kuwait and the restoration of the constitutional government of Kuwait. It is indeed heartening to see the world react swiftly and forcefully to check this unprovoked aggression. The damage and suffering brought about by this aggression are huge, both materially and psychologically. Kuwait, a small, peaceful and thriving state, albeit with its own share of problems caused by accelerated growth, has been turned into a shambles. The lives of the Kuwaitis and other nationals residing in Kuwait have been abruptly disrupted, and more than half the population have fled in fear. We have all seen on the television screen the plight of the Asian workers who fled Kuwait, only to be trapped in the makeshift refugee camps on the borders of Iraq and Jordan. The few video pictures and the eye-witness reports from those lucky ones who were able to escape the forced guest status in Kuwait tell of a sad wholesale destruction of beautiful and peaceful Kuwait.

Lost in the news are the stories being told of the mad and irrational dismantling and destruction of scientific and research institutions in Kuwait. In particular, eye-witness reports tell of systematic and organized theft from the Kuwait Institute for Scientific Research, KISR. Computers, scientific instruments, books, supplies and even the rugs have been carted off. One of the leading scientific research institutions of the Arab world is now in disarray. Important research experiments in biology, ecology and the environment have been interrupted and, in some cases, years of field and laboratory work will have to be repeated. KISR was a truly international research institution; it employed more than 1,000 scientists, engineers and support staff from about 40 different nationalities, including those from the United States and Canada, Western and Eastern Europe, Asia, Africa and the Arab world.

KISR research programmes cover a wide range of mission-orientated research areas related to Kuwait, the Arab world and the developing countries. Among the major research activities pursued at KISR are: sea-water desalination and natural water resources, solar and other forms of renewable energy, fisheries and oceanography, biotechnology, arid land agriculture, air pollution and other desert and marine environmental problems of the Gulf region, petroleum and petrochemical technology, urban engineering, computer and microprocessor applications, applied economic and operation research

and other practical research.

KISR's new facilities, completed in 1986, include some of the most modern multipurpose applied scientific laboratories in the world, a major scientific information centre with over 100,000 scientific books and periodicals supported by automated information systems connected to data centres elsewhere. The KISR computer centre features many mainframes, including a newly acquired IBM 3090 and hundreds of PCs and workstations. KISR is empty now. We are told that the buildings are being used as military headquarters. Other scientific and educational institutions in Kuwait such as Kuwait University have fared no better.

All present and previous colleagues at KISR and other scientists who visited KISR will be deeply saddened at what has happened to KISR, and will join in calling upon our colleagues throughout the world to voice their opposition to such shocking actions, and hope that the world scientific community will call upon the Iraqi authorities to refrain from such destructive action and to restore KISR properties and facilities. We are confident that, as soon as Kuwait regains its independence and sovereignty, the world scientific community will help KISR to become once again a centre for creative applied research for the benefit of all mankind, the vision and the dream that guided all those who helped build KISR and make it a reality.

ADNAN A. SHIHAB-ELDIN
(Principal Research Scientist and former
Director General, KISR)
*Lawrence Berkeley Laboratory,
1 Cyclotron Road,
Berkeley, California 94720, USA*

RICHARD WILSON
*Department of Physics,
Harvard University,
Cambridge, Massachusetts 02138, USA*

SIR—I must take issue with your leading article "Beyond the trouble in the Gulf" (*Nature* 346, 683; 1990). You suggest that Iraq was the only Middle East oil-producing country that has spent oil revenues on a national programme "... being militarily strong in this case".

The leaders of the United Arab Emirates (UAE) have invested heavily in national programmes, including health, education, conservation, the greening-of-the-desert and agriculture. The leading article also suggested that "... the oil producers of the Middle East ... seem not to have turned their minds to creating a first-rate university somewhere". This statement does a disservice to the long-standing efforts of the leaders of the Emirates, particularly H. H. Sheikh Zayed Bin Sultan Al Nahyan, president of

the UAE and H. H. Sheikh Nahyan Bin Mubarak Al Nahyan, chancellor of the United Arab Emirates University.

They and the dedicated faculty, drawn from other Arab countries as well as from Western countries, are working very hard to create an outstanding, modern university at Al Ain, Abu Dhabi. Those of us working with the United Arab Emirates University are impressed with the commitment of funds and the insistence on excellence that characterize their effort. In conjunction with the UAE University, H. H. Sheikh Zayed Bin Sultan Al Nahyan has established a new Desert and Marine Environment Research Center, dedicated to conservation and exploitation of the natural resources of the UAE. The Emirates have also invested in a strong system of technical colleges covering civil engineering, electronics, business and computer skills.

SIDNEY B. SIMPSON JR
*University of Illinois,
Department of Biological
Sciences (M/C 066),
Box 4348, Chicago, Illinois 606030, USA*

Doctor who?

SIR—Many of us must wonder what proportion of the references that adorn scientific papers have actually been looked up, let alone read and inwardly digested. An example has recently come to my attention of the creation of an entirely new scientific personality as a result of casual citation. A number of articles concerned with fibroblast growth factor have referred in the past few years to the first observation of mitogenic activity in brain extract as having been made by Trowell, Chir and Willmer in 1939¹. It is probable that the middle author made only a limited contribution to this work as he is actually the surgical qualification held by the first author. In those days, *Journal of Experimental Biology* listed the degrees of the authors, and so their names on the title page read "O. A. Trowell, M.B., B.Chir. and E. N. Willmer, M.A." The neogenesis of Dr Chir is thus easily explained.

Students of the history of embryology might be interested to know that it was also in 1939 that the first description was published of a mesoderm-inducing activity from adult tissues². Had the two groups got together, they might have anticipated by nearly 50 years the modern flurry of activity on the role of growth factors as agents of embryonic induction.

J. M. N. SLACK
*Imperial Cancer Research Fund
Laboratories,
Department of Zoology,
South Parks Road, Oxford OX1 3PS, UK*

1. Trowell, O. A. & Willmer, E. N. *J. exp. Biol.* **16**, 60–70 (1939).

2. Chuang, H. H. *Wilhelm Roux' Arch. EntwMech. Orgs.* **139**, 556–638 (1939).

Order in the midst of chaos

Complex systems, as different as neural networks and clusters of interacting galaxies, are susceptible to chaotic behaviour. Numerical calculations now suggest that some degree of order may persist beneath the chaos.

WHY is the study of complex systems offered as a means of solving all the outstanding problems of the real world? The explanation is not all that deeply hidden. The complex problems of the real world may have more in common with each other than with any of the simple problems dealt with elegantly in the textbooks. A neural network, for example, is a large number of neurons of which all interact with all others, so that people are reduced either to making generalizations about their expected behaviour or to making numerical calculations of particular systems. Globular clusters, or clusters of interacting galaxies, are in much the same case, as are the problems of fluid dynamics in the real world. All these systems, for example, share a propensity to run to chaos.

This is the spirit in which Kunihiro Kaneko from the University of Tokyo (and, in particular, from the teaching-orientated College of Arts and Sciences in which undergraduates spend their first two years) has managed rigorously to demonstrate a general property of systems of this kind (*Phys. Rev. Lett.* **65**, 1391; 1990). He says that neural networks and fluid dynamics are among the problems to which his calculations are relevant. But his result is interesting and important — that order can persist even when a complex system has lapsed into chaos.

Those wishing to follow Kaneko's argument had better start with a paper published a year ago (*Phys. Rev. Lett.* **63**, 219; 1989) which, among other things, contains both the notation for handling the problem and the inspiration for the study now described. The most arresting of his conclusions, then, was a kind of phase diagram outlining the relationship between chaotic states of a complex system with global coupling in which the phases are characterized by differing degrees of persistent order.

Formally, there is no great difficulty in setting up the problem. If there are N elements in a complex system, the time evolution of a property x_i of the i th element will be determined primarily by that element's autonomous evolution, but also by its interactions with all the other elements in the system. So, if time is measured as a multiple n of some unit interval, the problem is simply that of relating $x_i(n+1)$ to the values of all the $x_k(n)$ s one time-step earlier. Formally again, $x_i(n+1)$ is predominantly

$(1-\varepsilon)f(x_i(n))$ on account of its autonomous evolution, and a sum of terms such as $(\varepsilon/N)f(x_k(n))$ on account of the interaction, where $f(x)$ is some function and ε is a number small compared with 1.

The fun begins with the choice of the function f , which, if linear, yields nothing remarkable, but which is well known to yield chaotic behaviour if taken to be, say, $f(x) = 1 - ax^2$ (where a is again a constant) with the rule that the integral part of any calculated x is discarded. The outcome, now one of many textbook examples of how to generate chaos, has the effect of mapping the range of numbers between -1 and $+1$ onto itself with each successive iteration, and also has the convenient property that one can regard positive and negative values of the variable x , as simply minuses and pluses, corresponding to, say, OFF and ON states of a neuron, according to whether they lie below or above a characteristic value of the parameter a .

Kaneko's results are not dependent on this choice of function; he has been able to show that the same general pattern of behaviour crops up with other functions. But the dynamical rule specifying the evolution of the system is rather special, in that all the interactions between different elements have the same form and the same strength. Those assumptions evidently oversimplify the properties of, say, a true neural network or globular cluster, but there is no reason to suppose that the assumption is critical to what emerges.

And what is that? The conclusion that order is curiously immune to the onset of chaos. What Kaneko showed last year is that coherent oscillation of a system from PLUS to MINUS is surprisingly common, but that appropriate values of the parameters ε and a cause it to break up into two clusters (defined conceptually, not geometrically) which oscillate out of phase with each other, members of one set going from PLUS to MINUS while the others go from MINUS to PLUS. These persistent states of chaotic oscillation are the 'phases' provided by the model. What seems to be happening is that, for each element, there is a competition between its tendency to chaotic instability and its tendency to conformity arising from the averaging effect of the system of elements as a whole. The greater the nonlinearity (parameter a), the more widespread the disorder; the greater the averaging effect (parameter ε), the more persistent is

coherence.

Kaneko's article last year raises a host of questions that have not yet been answered, but his new calculation answers one of them in a fashion that raises further intriguing questions. However remote the model may seem from the models of classical dynamics, it does at least provide an easy way of going over to a mean-field representation of the system in the Landau and Ginzburg sense. Indeed, the quantity $(1/N)\sum f(x_k(n))$ represents a kind of ordering mean field whose statistical properties can be calculated numerically. What emerges is that, for each set of parameters, the statistical distribution of the mean field is an inverted bell-shape centred on some numerical value and that there is a lower limit to the width of the inverted bell (or to the variance of the mean field) with increasing complexity (the value of N). That is what would be expected from the central limit theorem of statistics.

The surprise is what happens to the mean-square deviation of the mean field, which would naively be expected to fall to zero with increasing complexity by what is still quaintly called the law of large numbers. The numerical experiments show that, for some values of the parameters, the mean-square deviation of the mean field falls off not to zero, but to some constant value which is, of necessity, greater than zero. What can be the explanation? Order in the midst of chaos. Kaneko shows this to be the case by calculating a simple statistic that measures the mutual correlation of the states of elements in his model of the complex system. For good measure, he shows that the expectations of the law of large numbers are fulfilled if a little noise is added to the system, but that reality does not approach expectation as $1/N$, but as some power of that factor less than one.

That chaotic behaviour does not necessarily imply that everything is random is borne out by the well-known regularities in the way in which chaos develops — for example, by bifurcation of period-doubling, for all the world as if the development of chaos is an orderly process. The recognition that the chaotic states of a system may embody a measure of persistent coherence may be even more significant if, say, it is a way of accounting for persistence of neural memory in disordered neural networks.

John Maddox

The flight of the bumblebee

Jeremy M. V. Rayner

BIOLOGISTS are wont to characterize the engineer by his belief that the bumblebee cannot fly. They have had to rely on the empirical observation that the bumblebee *does* fly, but up until now have had little insight into its flight mechanics and physiology. On page 472 of this issue¹ Charles

measurement of the tiny changes in gas concentrations produced by an insect in forward flight, and the only results comparable with those for vertebrates were for a few tethered insects¹. Ironically it is the bumblebee paradox that has made the new results possible: for its size the energy requirements of flight are sufficiently large that — with some ingenuity — they can be measured. By designing a special sealed wind-tunnel which uses a rubber bladder to allow tunnel volume to compensate for pressure changes, and adapting an oxygen analyser to enhance its sensitivity, Ellington and colleagues were able to quantify the rate of change of oxygen concentration in the tunnel, and from this to determine the bee's power output at different speeds.

It is natural to ask how far these results can be generalized to bumblebees outside the wind tunnel, to flight of other insects or to flight performance in all animals. The wind tunnel is an artificial environment, and may distort flight behaviour and energetics (this is certainly the case for some of the bird measurements²). The results obtained by Ellington *et al.* do not indicate any obvious problem: the measured oxygen consumption of 50–60 ml O₂ g⁻¹ h⁻¹ for a 500 mg bumblebee is equivalent to a power output of 0.14–0.17 W, comfortably in agreement with the value of 0.17 W obtained by extrapolation from the available vertebrate measurements³ (though the precision of extrapolation to such a small size is low). Broadly similar results are obtained by modelling the aerodynamics of the bumblebee wingbeat⁴, and it seems reasonable to suppose that these measurements of flight energetics are a good estimate of flight costs of bees or other Hymenoptera in the wild⁵.

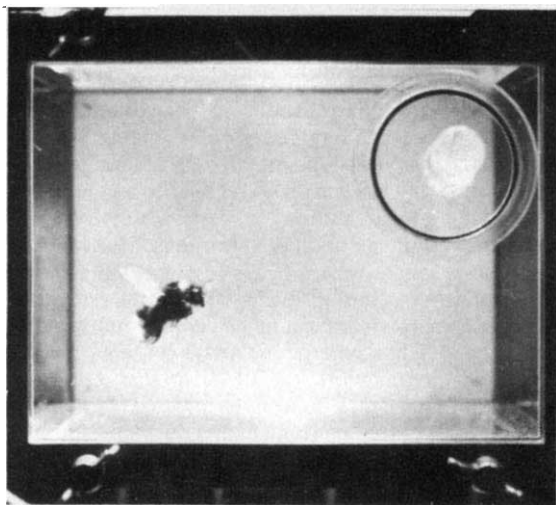
The wind-tunnel technique is costly and time consuming, and generalization to other animals relies on theoretical modelling. The most remarkable conclusion of the new results is that energy output is independent of speed from hovering up to 4 m s⁻¹ (that is, approaching the maximum speed used by

wild bumblebees). This runs counter to the predictions of existing aerodynamic theories^{2–5,7,8}. The main factors determining power output are the aerodynamic drag forces on the wings: the different components of drag change magnitude as speed varies, and the curve of mechanical power output (the rate of increase of kinetic energy of the air surrounding the insect) against speed is U-shaped, falling and then rising as the animal flies faster: power is particularly large in hovering flight^{2–5,7,8}. The models are best developed for birds and bats, but much the same shape of curve applies to fixed-wing aircraft and helicopters, and the underlying arguments are equally valid for insects. The latest model of bumblebee flapping flight⁵, developed alongside the new measurements, shows a flatter U-shaped curve than previously, but still fails to explain why forward flight is not cheaper than hovering.

We must, of course, question the extent to which we should expect theory and experiment to agree. Aerodynamic models estimate only part of the total energy cost, for the animal also consumes energy in converting fuel to mechanical and muscular work, and in internal processes such as digestion, and fluid and gas transport. The ratio between mechanical output and metabolic input energy rates is quantified as efficiency, which encompasses a range of different factors from flight muscles to respiration. For most flying animals efficiency lies in the range 10–25 per cent, and depends on the animal's size^{3,4,9}; estimates of efficiency in living animals are obtained by comparison of aerodynamic theory and physiological experiment, and provide no opportunity to identify possible flaws in either approach.

One resolution is that efficiency varies with speed in such a way that total power is approximately constant, but this does not seem to be an adequate explanation^{4,9} and it enjoys no theoretical justification. In the absence of a means of determining effi-

C. Ellington

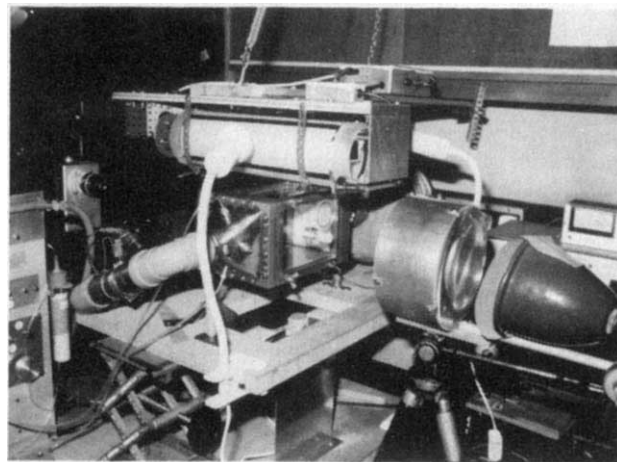


Test flight — measuring the oxygen consumption of a bumblebee flying in a wind tunnel.

Ellington and the late Ken Machin of the University of Cambridge, and Tim Casey of Rutgers University, report the remarkable achievement of measuring the oxygen consumption of flying bumblebees. These results provide the first direct measurements of the dependence of the energy cost on flight speed for a free-flying insect.

The claim of the 'impossible' bee is embedded in folklore, and a search of the literature fails to uncover its original source. The argument is that the bee's wings are so small that the power it would need to fly greatly exceeds that which its muscles can produce; moreover, even if it could create enough power it would have to fly impossibly fast to obtain sufficient lift. No self-respecting aerodynamicist would design an airframe of such proportions. But flight is a central feature of the biology of the bumblebee, and we must suppose that its superficially unfavourable design is actually appropriate for its lifestyle.

Measurements of gas exchange (and hence flight energetics) as a function of flight speed are familiar for about ten species of birds and bats^{1–4}. All the studies have employed wind tunnels because flight speed cannot otherwise be controlled: measurements of hovering and of unrestrained free flight cannot be generalized, and are hard to compare with existing aerodynamic theories. Hitherto, technical difficulties have prevented accurate



C. Ellington

The special wind-tunnel used to determine the oxygen consumption of a flying bee. A rubber bladder allows tunnel volume to compensate for changes in pressure.

ciency, measured metabolic energy gives a much closer indication of the behavioural and ecological costs of flight than can aerodynamic models. Because similar flat power-speed curves have been found in some birds¹ (although not in bats) it is advisable to avoid making incautious use of theoretical estimates of flight costs: the approach² of assuming a constant efficiency independent of speed and size cannot at present be justified. Reconciliation of physiological and aerodynamic approaches to energetics, and in particular the deeper understanding of efficiency, remains the main challenge facing students of animal flight.

Finally, why is it that the bumblebee has such small wings? The answer must lie in the bee's flight ecology, that is in the way it uses flight to gather nectar and pollen. This can lead it to carry large loads, for which small wings are not especially suited^{8,10}. Hummingbirds (and also some species of bats) also feed on nectar, often hovering by host plants, and also have

smaller wings than expected for their size^{8,10}. In these vertebrates the explanation seems to be the need to fly fast to defend large territories and to minimize time in commuting flights; the same explanations seem equally plausible for bumblebees. □

Jeremy M.V. Rayner is in the Department of Zoology, University of Bristol, Woodland Road, Bristol BS8 1UG, UK.

1. Ellington, C. P., Machin, K. E. & Casey, T. M. *Nature* **347**, 472–473 (1990).
2. Pennycuik, C. J. *Bird Flight Performance: A Practical Calculation Manual* (Oxford University Press, 1989).
3. Norberg, U.M. *Vertebrate Flight* (Springer, Heidelberg, 1990).
4. Rayner, J. M. V. in *Physiological and Ecological Aspects of Bird Migration* (ed. Gwinner, E.) (Springer, Heidelberg, in the press).
5. Dudley, R. & Ellington, C. P. *J. exp. Biol.* **148**, 53–88 (1990).
6. Wolf, T. J., Schmid-Hempel, P., Ellington, C. P. & Stevenson, R. D. *Funct. Ecol.* **3**, 417–424 (1989).
7. Gnosspelius, O. F. *Jl R. aeronaut. Soc.* **29**, 543–547; 648–649 (1925).
8. Rayner, J. M. V. *Curr. Orn.* **5**, 1–77 (1988).
9. Ellington, C. P. *J. exp. Biol.* **115**, 293–304 (1985).
10. Norberg, U. M. & Rayner, J. M. V. *Phil. Trans. R. Soc. Lond.* **B316**, 335–427 (1987).

MATERIALS TECHNOLOGY

The power of paradox

Robert W. Cahn

SCIENCE progresses by the resolution of paradoxes. Interpretations of the nature of light swung back and forth — particles or waves? — until the quantum synthesis embraced both absolutes. A door must be either open or shut, the logical French tell us. But a liquid crystal is neither solid nor liquid; a quasicrystal, neither crystalline nor amorphous; a magnetic 'fluid' can be either mobile or rigid. The best doors are both open and shut.

So it is with strength. Ceramics are much stronger than metals: but drop a plate on the floor and it shatters, drop an aluminium bottle and it dents and survives. What really matters is toughness, the ability to resist shock, to break gradually rather than catastrophically, to give adequate warning of impending disaster. Brute strength will not suffice. The introduction of fibre-reinforced composites — arrays of strong ceramic fibres in weak polymeric matrices — as a practical way of taming the brittleness endemic to ceramics expressed the principle of "strength made perfect in weakness", to cite an apposite biblical paradox. The fibres arrest cracks which try to spread through the polymer, and also carry most of the load. But polymers will not stand up to red heat.

Now, as reported elsewhere in this issue (W. J. Clegg, K. Kendall, N. McN. Alford, T. W. Button, & J. D. Birchall, *Nature* **347**, 445–447; 1990), a group from the ICI Advanced Materials Laboratory has found a new way of creating toughness by marrying strength and weakness; their

way produces a composite material which is relatively cheap to make and moreover contains only components resistant to high temperatures. The combination of intrinsic ceramic strength and heat resistance with both toughness and cheapness is something new.

Clegg and his collaborators have prepared blocks consisting of silicon carbide layers, separated by thin graphite interlayers. Graphite in itself is far from weak — in the form of carbon fibres, it is one of the strongest materials known — but here it forms a weak interface: a crack in the SiC impinging upon the interface becomes diverted so that it carries on spreading along the interface, dissipating energy as it does so. Silicon carbide and graphite are both strong but the interface between them is weak in shear, not in tension or compression: strength made perfect in weakness. The material, in fact, behaves very much like nature's own favourite composite, wood.

The composite is made by mixing silicon carbide powder, doped to aid sintering, with a concentrated polymer solution in a special high-shear processor. The ICI team had previously developed both the solution and the mixer to make ultra-strong cement. The resulting SiC paste, in the form of 2-mm sheets, is rolled out like dough into layers 0.2 mm thick, coated with graphite, stacked in tenfold multilayers, slowly dried out and sintered. Graphite, we are told, has the crucial property of not reacting with stoichiometric SiC, which can dissolve no excess of either

of its constituents. For comparison, the authors also sintered unrolled, uncoated SiC sheets.

Slow bend tests at ambient temperature show, for the composite, gradual fracture with repeated load drops as incipient transverse cracks on the tension side are diverted to spread laterally along SiC/graphite interfaces. A specific work of fracture — one test of toughness — ranging from 4,600–6,700 J m⁻² was recorded, compared with a value of only 60 J m⁻² for the monolithic (non-composite) material (a figure further reduced by a technical correction to 30 J m⁻²). The composite is thus more than a hundred times tougher than the monolith, at least in bending and at room temperature. The stress at which failure begins in the composite is reduced by 25 per cent, however: every benefit has its price.

The work of fracture of the composite is slightly higher than that of the hardwood, deal. So the ICI team has improved upon nature in two ways: their material is tougher than hardwood and it can certainly be safely heated to red heat. Whether the material maintains its strength and toughness at high temperature we are not told.

The new material is, in effect, a ceramic/ceramic composite. Others such have been made before, but only by expensive methods such as gas-phase reactions which take months to make what the ICI method produces in hours. Carbon/carbon composites (S. E. Hsu & C. I. Chen in *Superalloys, Supercomposites and Superceramics* (eds J. K. Tien & T. Caulfield) 721–744; Academic, New York, 1989) are an exception: they are an industrial product category in actual use, and are made by impregnating woven carbon fibre structures with liquid or gaseous polymeric or other organic precursors. However, they are mostly used for applications such as rocket nozzles or fighter wheel-brakes for which cost is scarcely considered. Very strong micro-layered aluminium/transition metal composites have been made at the Royal Aerospace Establishment in Farnborough by vapour deposition: this kind of alloy is very strong and tough but seems for the time being to have faded from view, perhaps because of its exorbitant cost.

Clegg *et al.* point out that their new technique is not restricted to making sheets. For instance, wires of SiC/carbon composites can be extruded from the paste and then coated and pressed together to make rods. But the real novelty of the approach is undoubtedly the probable cheapness of the processing method. □

Robert W. Cahn is in the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.

Mimics — or gimmicks?

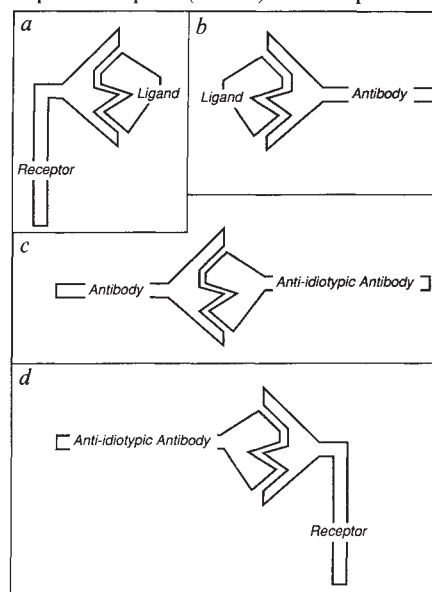
David I. Meyer

The principle is simple yet elegant. An anti-receptor antibody can be produced without having to use the receptor as an antigen. The technique is based on the fact that the binding both of a ligand to its receptor and of an antibody to an antigen depends on protein-protein interactions involving subtle conformational complementarity. A first antibody is raised that recognizes a ligand in a fashion that mimics recognition by the physiological receptor. The combining site of this receptor-mimicking antibody has structural features similar to those of the receptor's own ligand-binding site. If a second set of antibodies is produced against the combining site, and is similar in structure to the regions of the ligand that bind to the physiological receptor (see figure). These second-generation, or anti-idiotypic (anti-IT) antibodies, mimic the original ligand in binding to its receptor, with the obvious advantage that affinities should be high enough to fish out that receptor.

On page 444 of this issue¹, Pain *et al.* describe studies in which the anti-idiotypic approach was used to identify a receptor for the import of nuclear-encoded proteins into mitochondria. The experiments are predicated on the assumption that import is mediated by membrane-bound receptors that recognize and bind the N-terminal signal sequences of the precursors of these proteins. By first producing an antiserum against a synthetic version of the signal sequence of cytochrome oxidase subunit IV, and then raising anti-idiotypic antibodies to the original antiserum (see figure), Pain, Murakami and Blobel have identified three proteins that could be factors recognizing mitochondrial precursors.

One of these, of relative molecular mass 32,000 (M_r , 32K), is an integral membrane protein and thus a strong candidate for the receptor. Biochemical and functional analyses carried out by the authors provide evidence for this role. As would be expected, the anti-(IT) antibody can compete with the signal sequence (ligand) for binding to the putative receptor, and immune complexes including precursor and receptor could be isolated from cell-free extracts. Moreover, the anti-IT antibody, as well as regular antibodies raised against the 32K receptor, inhibited import of protein precursors in an *in vitro* assay. Morphological studies localized the receptor to zones of contact between inner and outer mitochondrial membranes. On page 488, a report from the same laboratory, by Murakami *et al.*², describes the results of the cloning, sequencing and disruption of the gene encoding the 32K receptor. Cells lacking the gene had severely disabled

mitochondria, and the import of some, but not all, proteins was significantly reduced. A search of protein-sequence data banks showed that this mitochondrial import receptor (MIR1) was 40 per cent



The anti-idiotypic technique. *a*, A receptor's binding site for its ligand is characterized by a conformation that ensures specificity. *b*, An antibody raised against the appropriate determinants on the surface of the ligand may have a structure in its combining site that mimics the ligand's receptor. *c*, A second set of antibodies, that recognize the antigen-combining site of the first antibody, are internal image anti-idiotypic antibodies whose own combining sites mimic the determinants on the ligand that mediate binding to its receptor. *d*, These anti-idiotypic antibodies can then be used to identify the receptor.

identical to mammalian phosphate transport proteins.

This brings to at least five the number of polypeptides that have been implicated in the import of proteins into mitochondria. Using antibodies raised against individual outer membrane proteins of *Neurospora*, two mitochondrial outer membrane (MOM) proteins of M_r 19K and 72K have been isolated^{3,4}; the yeast receptor MAS70, homologous to MOM72, as well as a 42K internalization site protein (ISP42), have been shown by genetic means to block import of the ATP/ADP carrier in the case of MAS70, and all precursors examined in the case of ISP42 (G. Schatz, personal communication); and using affinity-based techniques, two other groups have isolated 29K proteins on the basis of their affinity for synthetic mitochondrial signal sequences^{5,6}.

By almost bizarre coincidence, the use of signal sequences or anti-idiotypic antibodies as bait on fishing expeditions has landed previously characterized transport

proteins. One of the 29K putative mitochondrial import receptors, isolated through its high affinity for the signal peptide of ornithine transcarbamylase, has turned out to be an ATP/ADP carrier (G. Shore, personal communication). Cloning and sequencing of a pea chloroplast import receptor, also identified using the anti-idiotypic approach, show it to be the homologue of the spinach chloroplast triose phosphate-3-phosphoglycerate-phosphate translocator⁷. The sequence of MIR1 (see page 490), with its 40 per cent identity to mammalian phosphate transporters, is indeed identical to a recently cloned and sequenced yeast protein⁸ characterized by its ability to transport phosphate into mitochondria. All of these transport proteins are believed to be located in the mitochondrial inner membrane.

With these findings in mind, is the notion of a bifunctional translocator acceptable? Can the same protein transport both phosphate and protein precursors? From recent evidence it would seem that a single mitochondrial protein can have two distinct functions. The *Neurospora* enhancing protein of the matrix-processing peptidase complex is also a component of the cytochrome reductase complex in the inner membrane⁹. At this stage there is no obvious overlap in the function of these two enzymic complexes. In the light of this precedent, and given that certain bacterial proteins have been shown to have dual roles, a model based on bifunctional receptors is an easy one to propose. If, on the other hand, one excludes this possibility, a closer, more critical look at the use of anti-idiotypes in pulling out intracellular receptors would be in order.

The most successful use of anti-IT antibodies has been in the identification of cell-surface receptors for hormones, neurotransmitters, neuropeptides and growth factors¹⁰. In each of these cases, the proof resides in the ability of the antibody to mimic the ligand by eliciting the same physiological response *in vivo*. In the case of import of proteins into mitochondria, the anti-idiotypic antibody would have to mimic the precursor and be translocated into the matrix. Although many macromolecules, including DNA, can be imported into mitochondria by piggy-back on signal sequence-bearing chimaeras¹¹, antibodies cannot¹². In lieu of this, indirect methods have been employed, with the following circular reasoning as the result. When the affinity of a ligand for its receptor is believed to be low, conventional binding techniques for purification are replaced by producing anti-idiotypic antibodies against the ligand. Under circumstances where no physiological response can be mimicked, ligand binding is used to constitute proof of receptor activity. If binding experiments are feasible for proving receptor

function, they should be feasible for receptor isolation.

The anti-idiotypic approach has recently been employed by Vaux *et al.*¹³ to isolate an intracellular receptor involved in the retention of proteins in the lumen of the endoplasmic reticulum (ER). As polypeptides retained in the ER all share the sequence KDEL at their C terminus, anti-idiotypic antibodies raised against the appropriate peptides led to the identification of an intracellular membrane glycoprotein of M_r 72K. The proof of its KDEL receptor activity was predicated on its ability to bind KDEL sequences *in vitro*.

The yeast homologue of the ER retention receptor recognizes a slightly different sequence (HDEL for *S. cerevisiae* and DDEL for *Kluyveromyces lactis*), and has been identified genetically^{14,15}. Analysis of retention-defective mutants in yeast resulted in the identification of the HDEL and DDEL receptors as the products of the gene *ERD2*. The yeast HDEL/DDEL receptors are smaller (M_r , 26K) and are not glycosylated. In a report soon to be published¹⁶, the same group has used the polymerase chain reaction to isolate the mammalian (human) homologue. Sixty-one per cent of the amino acids of the human homologue are identical in the sequences of one or the other yeast receptors, a very high degree of conservation. Although this poses a paradox that will have to be resolved, it is clear that the anti-idiotypic approach did not pick up the product of the mammalian equivalent of *ERD2*. In the case of mitochondria, the anti-IT antibodies that detected MIR1 failed to react with the other protein precursor receptors MOM19 or ISP42.

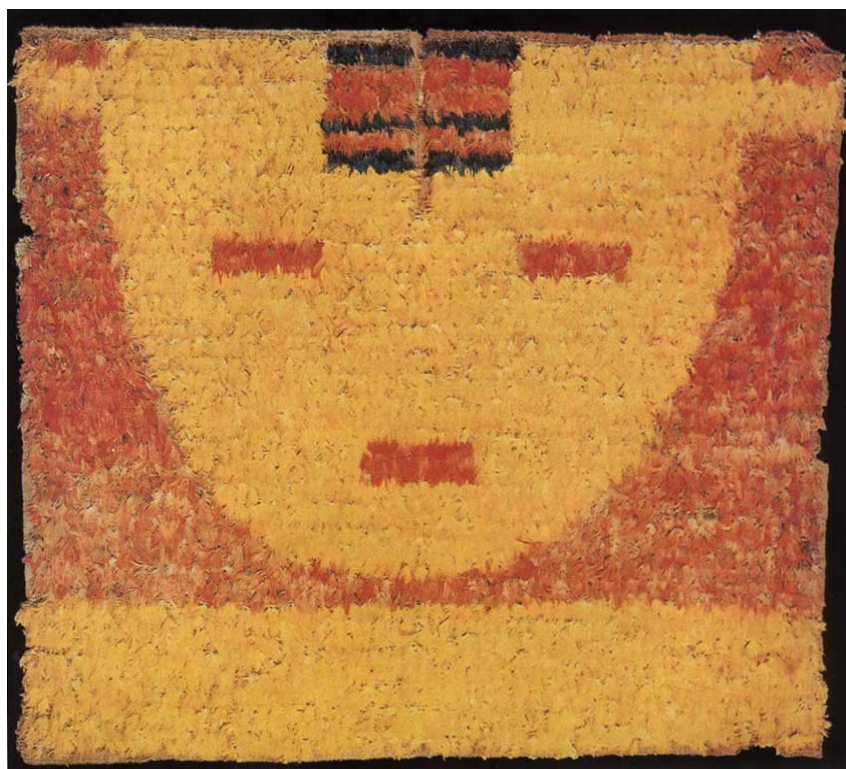
A feature of all of the studies using anti-idiotypic antibodies is that the isolated receptor binds specifically to the requisite ligand. This is proof-positive that the method works. Yet the nagging question persists: is it possible that out of the myriad of possible surface epitopes on the hundreds of different cellular proteins in a

given organelle, one or more just happens to be similar enough to a receptor-binding site to be picked up by this approach? These could include, for example, amphipathic, channel-forming helices of phosphate or nucleotide transporters. Research into protein transport, secretion and organelle assembly is currently dominated by the notion that protein precursors interact with moderate affinity with a host of intracellular proteins known as molecular chaperones¹⁷. An assumption is that the signal peptide is at least in part responsible for this interaction. Chaperones are one set of proteins that could be expected to present other protein precursor receptors and transporters with stiff competition for anti-idiotypic antibody binding, if for no other reason than their abundance. It may also be worth noting that phosphate and nucleotide transporter proteins are very abundant in mitochondria.

Given these limitations, it seems that the best use of the anti-idiotypic approach would be as a coarse screen for putative receptors, augmented with reverse genetics to verify function *in vivo*. The comprehensive approach taken by Murakami *et al.*² included reverse genetics to confirm that the protein extracted with the anti-IT antibody has receptor activity. In the case of the knock-out of the *MIR1* gene, the result was a phenotype consistent with either aberrant transport of phosphate or import of protein precursors. In the case of the KDEL receptor, one hopes that its homologue will soon be found in a system that can be genetically manipulated. Loss of retention of polypeptides in the lumen of the ER should be easily detected. □

David I. Meyer is in the Department of Biological Chemistry, School of Medicine, University of California, Los Angeles, California 90024, USA.

Art for the anthropologist



THIS half-segment of a ceremonial Nazca tabard from the south coast of Peru appears in an exhibition and sale currently running in London. Made by sewing macaw feathers onto a cotton base, the vivid colours owe their preservation to burial in sand. During the period 200 BC–800 AD, the Nazca culture produced arguably the most exciting feather textiles in ancient Peru: a common theme is the Sun, resplendent with rays and facial features. This example shows the Sun in more contemplative mood. The image was produced very simply but is very effective.

The tabard will be on show with other Andean feather textiles and artefacts until 2 November at Thomas Gibson Fine Art, 44 Old Bond Street, London W1X 3AF, UK. Prices start at around £4,000. Thomas Gibson normally deals in modern art — but the gallery, like artists themselves, recognizes the connection between so-called 'primitive' art and the motives behind modernist movements as diverse as surrealism and abstract expressionism.

H.G.

- Pain, D., Murakami, H. & Blobel, G. *Nature* **347**, 444–449 (1990).
- Murakami, H., Blobel, G. & Pain, D. *Nature* **347**, 488–491 (1990).
- Söllner, T., Griffiths, G., Pfaller, R., Pfanner, N. & Neupert, W. *Cell* **59**, 1061–1070 (1989).
- Söllner, T., Pfaller, R., Griffiths, G., Pfanner, N. & Neupert, W. *Cell* **62**, 107–115 (1990).
- Gillespie, L.L. *J. Biol. Chem.* **262**, 7939–7942 (1987).
- Ono, H. & Tsuboi, S. *J. Biochem. (Tokyo)* **107**, 840–845 (1990).
- Schnell, D., Blobel, G. & Pain, D. *J. Cell Biol.* (in the press).
- Phelps, A., Schobert, C. & Wohlrab, H. *Biochemistry* (in the press).
- Schulte, U. *et al.* *Nature* **339**, 147–149 (1989).
- Köhler, H. *et al.* *Meth. Enzymol.* **178**, 3–35 (1989).
- Vestweber, D. & Schatz, G. *Nature* **338**, 170–172 (1989).
- Schleyer, M. & Neupert, W. *Cell* **43**, 339–350 (1985).
- Vaux, D., Tooze, J. & Fuller, S. *Nature* **345**, 495–502 (1990).
- Semenza, J.C., Hardwick, K.G., Dean, N. & Pelham, H.R.B. *Cell* **61**, 1349–1357 (1990).
- Lewis, M.J., Sweet, D.J. & Pelham, H.R.B. *Cell* **61**, 1359–1363 (1990).
- Lewis, M.J. & Pelham, H.R.B. *Nature* (in the press).
- Hemmingsen, S.M. *et al.* *Nature* **333**, 330–334 (1988).

Pigments of the imagination

Paul Bahn

THE first direct dating of Ice Age cave-art in Europe has been carried out as part of a French programme of pigment analysis in the Quercy region of France¹ and the French Pyrenees^{2,3}. These results follow the recent discovery of human blood in red pigments in some Australian caves and rock-shelters, which was dated by radiocarbon to 10,730 and 20,320 years before present (BP)⁴, and show that the study of cave art has entered a new and highly promising phase of research.

As in Australia, it was the discovery of organic material in pigment used in French caves which produced the radiocarbon date. Analyses of pigments have consistently revealed the use of iron oxide (red) and manganese dioxide (black). But the improved analytical methods of recent years have produced far more detailed results, most notably from Lascaux cave⁵, and some years ago it was discovered that some of the black figures in the Pyrenean cave of Niaux contained charcoal⁶.

The research programme of pigment analysis in the decorated caves of Quercy⁷ revealed that charcoal was used here too, in the cave of Cougnac. It is a sample of 100 mg from a black dot on the cave wall, between a megaloceros (giant deer) figure done in charcoal and a 'speared human' figure, that has produced a radiocarbon date of $14,300 \pm 180$ years BP¹.

Samples of pigment from red figures on Cougnac's walls were compared with a deposit of red ochre on the cave floor, and with ochre sources outside the cave and 15 km away; the floor deposit was probably related to production of the red figures, and the ochre used was most likely obtained from local clays. Two points to emerge are that the large red animal figures in the centre of the cave's main panel were drawn with the same pigment and were therefore probably a composition, done in a single production event. And whereas stylistic studies had assigned Cougnac to the Early Magdalenian period, the date (together with one of 15,000 years BP from a reindeer bone in the cave) points to the Middle Magdalenian. Further dates, however, will be needed for confirmation.

Similar results have emerged from pigment analyses in the Pyrenees^{2,3}, especially in the caves of Niaux and Réseau Clastres. Dates have not yet been obtained from the charcoal figures at Niaux, but analyses of samples, carried out by scanning electron microscopy, X-ray diffraction and proton-induced X-ray emission, have revealed four specific 'recipes' of pigments mixed with mineral 'extenders' and binders: talc; baryte with potassium feldspar; potassium feldspar; and potassium feldspar with biotite (a

kind of black mica). Beyond making paint go further, extenders have other advantages; for example, adding biotite makes red paint spread easily when wet, produces a darker colour than pure red ochre, improves adhesion to the wall and stops the paint cracking as it dries.

In Niaux's famous 'Salon Noir', most of the animal figures were first sketched out in charcoal, and then manganese paint using recipe 4 was added on top. It therefore seems that the Salon Noir was indeed a 'sanctuary', a special place where the figures were carefully planned ahead of execution. The figures in all other parts of this huge cave were done more spontaneously, without preliminary sketches.

So these analyses are helping to establish how the caves were decorated: the Salon Noir contains all four recipes, though most of it was done with number 4, as was a sign at the far end of the cave; the Réseau Clastres, on the other hand, had only recipe 4, without preliminary sketches, confirming that it was visited very briefly. Another important application is helping in the detection of fakes: only one painting in Niaux, a red-painted fissure interpreted as a vulva, proved to have no extender. As the figure was not mentioned by the first scholars to study the cave, and there are initials nearby, it is clearly modern.

It is possible to 'date' recipe 4, because the occupation site of La Vache, directly across the valley from Niaux, has that same extender used with red and black paint on bones from layers dating to 12,850–11,650 years BP⁸, the Upper

Magdalenian; Niaux had traditionally been assigned to the Middle Magdalenian on the basis of style. Recipe 2 does seem to belong to the Middle Magdalenian, as it has been found on a bead from the cave of Enlène in levels of 13,940–12,900 years BP, and also in the Middle Magdalenian of the cave of Le Mas d'Azil.

One difference between the two projects is that the Quercy team believe the composition of its pigments to be entirely natural, because the mineral components discovered exist in the same proportions in local sediments. The Pyrenean team, on the other hand, insist that their compositions and recipes contain associations of components which it is quite impossible to find in nature. These claims are not necessarily contradictory — artists in Quercy may have been content to use natural pigments, whereas those in the Pyrenees may have concocted new ones.

Pigment analysis is providing a tool, applicable at least regionally, for dating Magdalenian paintings, and is revealing that behind apparent stylistic homogeneity there is technological heterogeneity. Styles lasted longer than was thought, and are of limited use in dating. □

Paul Bahn is a freelance writer on archaeology based at 428 Anlaby Road, Hull HU3 6QP, UK.

1. Lorblanchet, M. *et al.* *Rock Art Res.* **7**, 4–20 (1990).
2. Clottes, J., Menu, M. & Walter, P. *Rock Art Res.* **7**, 21–26 (1990).
3. Clottes, J., Menu, M. & Walter, P. *Bull. Soc. Préhist. française* (in the press).
4. Loy, T.H. *et al.* *Antiquity* **64**, 110–116 (1990).
5. Ballet, O. *et al.* in *Lascaux Inconnu* (eds Leroi-Gourhan, A. & Allain, J.) 171–174 (CNRS, Paris, 1979).
6. Brunet, J., Callède, B. & Orial, G. *Stud. Conserv.* **27**, 173–179 (1982).
7. Lorblanchet, M., Labeau, M. & Vernet, J.L. *Préhist. Quercinoise* **3**, 79–94 (1988).
8. Buisson, D., Menu, M., Pinçon, G. & Walter, P. *Bull. Soc. Préhist. française* **86**, 183–191 (1989).

BLACK HOLES

Making a compelling case

Charles D. Bailyn

BLACK holes, although perhaps the most spectacular theoretical constructs in astrophysics, are by their nature elusive to observers. They can be detected only indirectly, by their gravitational influence on their surroundings. Currently the most convincing black hole candidate is one member of a binary star system known to X-ray astronomers as A0620–00. New work by Haswell and Shafter¹ adds yet another brick to what has become an impressive edifice of data demonstrating that this system must contain a black hole.

Strong compact X-ray emitting objects are good places to start looking for black holes. The energy requirements for producing large numbers of X-rays in a small volume are severe. The only viable energy source for such emission is the gravitational energy released by material accreting

onto a compact object. Models along these lines have been used to account for quasars (sufficiently far away from us to require galaxy-sized black holes to account for their apparent luminosity) and for compact X-ray sources in our Galaxy. In a few cases, detailed studies of the latter systems have produced virtually airtight evidence for the existence in nature of black holes.

The strongest galactic X-ray sources are binary stars bound sufficiently closely for one member of the system to accrete material from the other. The accreting gas typically forms a disk around the accreting star, and becomes hotter as it spirals inwards. The immense X-ray emission observed from these systems implies that the accreting star must have an enormous gravitational field, and hence must be

extraordinarily dense. Only black holes, or neutron stars, in which the protons and electrons in ordinary matter are crushed together to produce an object with densities as high as several hundred million tons per cubic centimetre, will suffice.

Although various indications suggest that most of the compact objects in X-ray binaries are neutron stars, there is a strong limit on the mass a neutron star can attain. Models of the equations of state likely to pertain inside neutron stars² as well as more general relativistic constraints³ show that neutron stars with masses greater than three times that of the Sun cannot exist; such objects would inexorably collapse to become black holes. The object of the game in galactic black-hole hunting is therefore to find a compact object in an X-ray binary whose mass exceeds three times the Sun's.

A0620-00 commanded the immediate attention of X-ray astronomers when it erupted in 1975, becoming for a short while the brightest object in the X-ray sky. The optical counterpart was readily identified as a nova star called V616 Mon, which had been seen to erupt once before in 1917. Observations following the object's return to its quiescent level led to the determination of the binary's orbital period as about 8 hours⁴. At the same time, the companion star was shown to have velocity variations of almost 1,000 kilometres per second during each orbit⁴. With this information, a simple application of Kepler's laws demonstrated that the mass of the compact object could be no lower than 3.18 ± 0.16 times the Sun's. Very reasonable assumptions regarding the nature of the secondary star and the origin of the orbital luminosity variations then sufficed to put the minimum mass of the compact object unambiguously over the three-solar-mass limit.

Now, Haswell and Shafter¹ have measured the velocity variations of the emission lines associated with the accretion disk around the compact object. Because they originate in a disk which has gas flowing at high speeds both towards and away from the observer, these spectral lines are extremely broad (owing to Doppler shifting). Indeed the breadth of the lines, which indicates the velocity of the innermost edge of the accretion disk, has itself been used as an indication that the compact object must be significantly more massive than three solar masses⁵. By using simple models of the complex structure of the disk emission lines, Haswell and Shafter have shown that the accretion disk, and presumably the compact object which it surrounds, have orbital variations of about one-tenth the amplitude of those of the comparison star. This factor of ten represents the mass ratio between the compact object and its companion.

This information increases the minimum mass of the compact object to

3.82 ± 0.24 times the Sun's. Thus unlike Cygnus X-1 and LMC X-3, the more familiar black-hole candidates usually invoked in textbooks and popular articles, the compact object in A0620-00 can be demonstrated to be above the three-solar-mass limit from the dynamics of the binary system alone.

In fact, it seems likely that the black hole is substantially more massive than that. If the companion has the mass usually associated with stars of its spectral type, then the compact object must be at least four times the solar mass⁵; the very plausible models of the accretion disk structure used by Johnston, Kulkarni and Oke⁵ lead to a similar limit. The final piece of information required for an accurate measurement of the mass of the compact object would be the inclination of the orbital plane to the line of sight. In principle this can be determined from the ellipsoidal variations, small changes in brightness caused by our changing angle of view during an orbital period.

PALAEOCLIMATOLOGY

Estimating atmospheric CO₂

N. J. Shackleton

ONE of the great surprises from studies of the ice-age climates was the discovery, made from bubbles of fossil air trapped in the great ice sheets¹, that there was less CO₂ in the atmosphere during glacial periods than in interglacials. As well as contributing to our understanding of climate change (especially important given current fears about the effect of anthropogenically generated greenhouse gases), the discovery offers a means for integrating studies of the different contributions of the ocean, atmosphere and biosphere. Now, on page 462 of this issue², Jasper and Hayes report a new approach to estimating the history of CO₂ partial pressure in the ocean surface waters. The method should extend the CO₂ record well beyond the start of the ice-core record. Also, it should enable us to elucidate the changes in the dynamic equilibrium of CO₂ between the atmosphere and the sea surface.

Previous attempts to explain natural variations in atmospheric CO₂ have been based on Broecker's discussion³ of the oceanic 'biological pump' whereby marine photosynthesis removes carbon from the surface waters of the ocean and hence holds atmospheric CO₂ below the level it would attain if the ocean were abiotic. Previous attempts to apply his arguments have involved measuring variations in the ¹³C content ($\delta^{13}\text{C}$) of planktonic and benthonic foraminifera from a low-latitude deep-sea sediment core. The history of changes in the $\delta^{13}\text{C}$ difference between warm surface waters (monitored by the

Preliminary measurements⁴ of these variations suggest a mass for the object of about seven times the solar mass. Several groups are currently observing these variations, but as yet the results do not appear to be totally self-consistent and repeatable (J.E. McClintock, personal communication). But even without this last piece of the puzzle, the case for a black hole in A0620-00 seems compelling; any alternative would be even more bizarre, and would probably result in severe damage to much of current astrophysical theory. □

Charles D. Bailyn is in the Department of Astronomy, Yale University, PO Box 6666, New Haven, Connecticut 06511, USA.

1. Haswell, C.A. & Shafter, A.W. *Astrophys. J.* **359**, L47-L50 (1990).
2. Arnett, W.D. & Bowers, R.L. *Astrophys. J. Suppl.* **33**, 415-436 (1977).
3. Shapiro, S.L. & Teukolsky, S.A. *Black Holes, White Dwarfs and Neutron Stars* 257-264 (Wiley, New York, 1983).
4. McClintock, J.E. & Remillard, R.A. *Astrophys. J.* **308**, 110-122 (1986).
5. Johnston, H.M., Kulkarni, S.R. & Oke, J.B. *Astrophys. J.* **345**, 492-497 (1989).

planktonics) and ocean deep water (monitored by the benthonics) can be interpreted as a history of changes in the efficacy of the biological pump and hence provides a record of variations in the equilibrium CO₂ partial pressure in ocean surface water¹. But the complexities of carbon isotope gradients within the ocean and of isotopic exchange with the atmosphere severely limit this approach.

Jasper and Hayes² instead estimate the amount of isotopic fractionation that occurs in the photosynthetic step in which marine algae use sunlight to incorporate dissolved CO₂. This fractionation depends on the concentration of dissolved CO₂. Jasper and Hayes' analysis requires separate estimates of $\delta^{13}\text{C}$ of dissolved CO₂ in the surface water, and $\delta^{13}\text{C}$ in the organic material formed by photosynthesis, as a function of geological age. The first estimate is the easy one to make: it is obtained by measuring $\delta^{13}\text{C}$ in the carbonate shells of planktonic foraminifera. The more difficult measurement is of $\delta^{13}\text{C}$ for the photosynthesized organic carbon. Jasper and Hayes measure this for selected organic molecules by combining mass spectrometry with gas chromatography. The molecules that they chose are C₃₇ alkadienones, which are specifically associated with the prymnesophyte algae (coccoliths).

Earlier work⁵ gave an approximate form for the relationship between the degree of photosynthetic fractionation and the concentration of CO₂ in sea water. Jasper and Hayes estimate the two

unknown constants in the relationship empirically by using the measured fractionation in about half of their fossil samples and calibrating this against the concentration of CO₂ estimated from the ice-core data¹ at equivalent ages.

Obviously, this approach precludes the authors from generating an independent record of atmospheric CO₂. And the poor quality of the earlier core record means that they cannot even usefully test their remaining estimates. This is a pity as it would be interesting to determine whether the rise in atmospheric CO₂ does in fact precede the most recent deglaciation as registered in the ¹⁸O isotope record in the same core. What is clear, however, is that the rise in estimated atmospheric CO₂ significantly precedes the rise in local sea-surface temperature as indicated either by the 'UK-37' method (in which temperature is estimated from the ratio of di-unsaturated to tri-unsaturated C₃₇ alkenones^{6,7}) or by the influx of the warm indicator *Globorotalia menardii* at the Y-Z faunal boundary⁸.

Much more work will be needed before this new approach actually refines our picture of the history of the atmosphere. However, the method will be very important for at least three reasons. First,

it could provide an atmospheric record directly linked to the marine stratigraphic records (perhaps offering an exciting direct way of correlating ice-core and marine records). Second, the limited age of the ice sheets means that they can yield a record extending back only a few hundred thousand years, whereas the new method can look back perhaps as much as 200 million years. Third, because the technique actually measures the oceanic partial pressure of CO₂, it may be possible to take slices from the record and to show for critical times (for example, the last glacial maximum) the sources and sinks of atmospheric CO₂ over the ocean surface. □

N. J. Shackleton is in the Godwin Laboratory for Quaternary Research, New Museums Site, Free School Lane, Cambridge CB2 3RS, UK.

1. Barnola, J.M., Raynaud, D., Korotkevich, Y.S. & Lorich, C. *Nature* **329**, 408–413 (1987).
2. Jasper, J.P. & Hayes, J.M. *Nature* **347**, 462–464 (1990).
3. Broecker, W.S. *Prog. Oceanogr.* **11**, 151–197 (1982).
4. Shackleton, N.J., Hall, M.A., Line, J. & Cag, S. *Nature* **306**, 319–322 (1983).
5. Popp, E.A., Takigiku, R., Hayes, J.M., Louda, J.W. & Baker, E.W. *Am. J. Sci.* **289**, 841–844 (1989).
6. Brassell, S.C., Eglinton, G., Marlowe, I.T., Pflaumann, U. & Sarinthein, M. *Nature* **320**, 129–133 (1986).
7. Jasper, J.P. & Gagosian, R.B. *Paleoceanography* **4**, 603–614 (1989).
8. Kohl, B. *Init. Rep. Deep Sea Drill. Proj.* **96**, 657–670 (1986).

COMPARATIVE ZOOLOGY

Clues to the origin of whales

André Wyss

No evolutionary transformation among a major group of mammals has been so dramatic as that undergone by whales. This transition from a terrestrial to a fully aquatic existence is remarkable both in terms of the degree of change encompassed, and in the number of biological systems influenced.

Long obscured by an incomplete and fragmentary fossil record, details of this transition are only now beginning to emerge, due in no small part to the extraordinary discoveries in Pakistan¹ and Africa² by Gingerich and co-workers.

The African find, representing new specimens of the middle Eocene (roughly 45 million years old) whale *Basilosaurus isis* from the well-known fossil concentra-

tions of Zeuglodon valley in the Fayum of north central Egypt, includes the first apparently functional hind limbs known for any whale (see Fig. 1). This new evidence offers tantalizing hints regarding the precise ancestry of the Cetacea — the order of mammals that comprises all whales — among archaic groups of land mammals. What could have been the function of these diminutive structures, which made up scarcely 3 per cent of the length of the entire animal?

Apart from minute vestiges of pelvic and proximal limb bones embedded in the musculature of the ventral body wall, and infrequent atavistic hindlimb rudiments, modern whales lack posterior appendages entirely. This is but one manifestation of

the highly transformed nature of living cetaceans which leaves little basis for comparison with other mammals and poses the greatest difficulty in attempting to discern the phylogenetic affinities of the group as a whole. So we have to rely on a notoriously unyielding fossil record to provide morphologies to close the gap between generalized early Tertiary terrestrial mammals and the highly derived diversity of extant whales.

The hind limb of *Basilosaurus* adds, therefore, critical new evidence to the small pool of available comparative anatomical information and is in apparent agreement with the early suggestion by Flower³ of a close cetacean–ungulate relationship. This idea has been endorsed more recently on the basis of dental and other grounds by Van Valen⁴ and by Szalay⁵, who argued for a more specific link between cetaceans and mesonychid condylarths (the latter a problematic, primarily early Tertiary assemblage of ungulates — see Fig. 2). Critical is the arrangement of the metatarsals, the third and fourth of which are enlarged in *Basilosaurus*.

One might quibble whether or not this condition represents the strict arrangement (termed paraxonic) seen among some mesonychids and artiodactyls (even-toed ungulates), in which the digits are symmetrically disposed about a plane between digits III and IV (digit II is essentially lacking in *Basilosaurus*), but at face value the interpretation seems correct. Early mesonychids, however, have feet with five well-developed digits⁶, suggesting that whales and artiodactyls might have evolved from within the mesonychids, rather than from a closely allied stock.

But molecular evidence does not support the kind of cetacean–artiodactyl relationship implied by the anatomical data⁷. Most immunodiffusion and protein-sequence studies suggest not only a close association between cetaceans and artiodactyls, but the inclusion of the former within the latter, for example an exclusive link of cetaceans and camelids⁷. Such suggestions are not supported by palaeontological evidence.

The *Basilosaurus* hind limb is also notable in that it permits, for the first time, the comparison of hind-limb structure between two different groups of marine mammal: cetaceans and pinnipeds (seals and sea-lions: hind limbs are unknown in sirenians — dugongs and sea-cows). The contrast between the two groups is striking, helping to confirm the highly distinctive pinniped pattern, present even in the earliest known representative⁸, as a diagnostic characteristic of seals and sea-lions. The classification of *Basilosaurus* offers an insight into the problems associated with the very poorly understood phylogenetic relationships among major

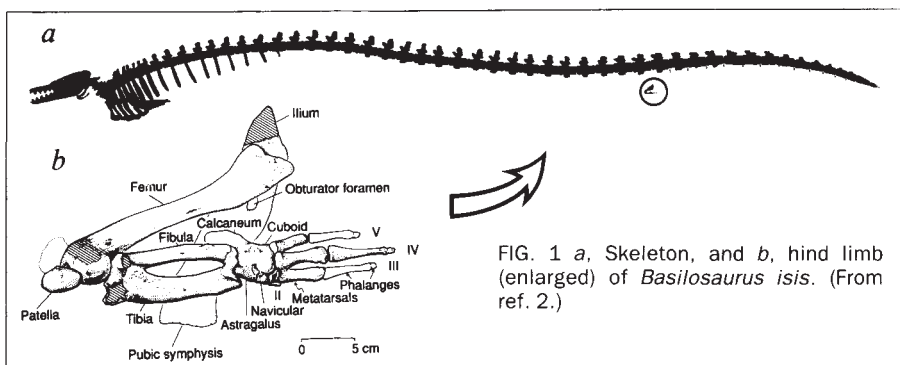


FIG. 1 a, Skeleton, and b, hind limb (enlarged) of *Basilosaurus isis*. (From ref. 2.)

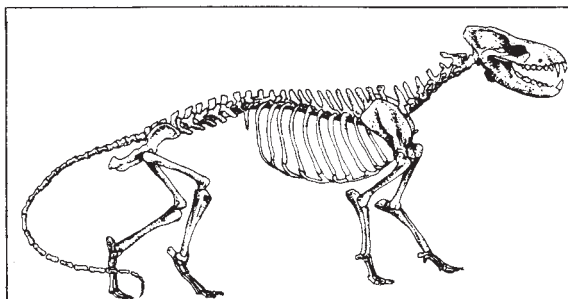


FIG. 2 Skeleton of *Mesonyx*, an early Tertiary ungulate believed to be close to the stock that gave rise to whales. (Courtesy R. J. G. Savage.)

lineages within the Cetacea. *Basilosaurus* is currently placed in the suborder Archaeoceti, a wholly extinct assemblage of early, toothed whales. As is widely recognized⁹, however, this group is defined solely by virtue of its containing Cetacea lacking the derived characters marking the two highly distinctive lineages of living cetaceans, mysticetes (the filter-feeding baleen whales) and odontocetes (the toothed echolocating whales). As such, 'Archaeoceti' (whales minus mysticetes and odontocetes) is a taxonomic waste-basket on a par with 'invertebrates' (metazoans that

are not vertebrates) and 'fish' (vertebrates without legs). Such unnatural taxonomic categories, unfortunately, are common throughout the Cetacea¹⁰ — scrapping them would be a valuable first step to an improved understanding of cetacean evolution.

Finally, as to the matter of what use *Basilosaurus* had for its tiny hind limbs, we can only guess. It is tempting to suggest that they might have served as guides during copulation², but one should keep in mind the difficulty of establishing whether features of organisms are adaptations towards particular functions, or if they are simply relics of an organism's evolutionary heritage that need not necessarily have any function at all¹¹. The problem of inferring function from structures is hard enough in living forms; structures that have no counterpart in the living world pose an even greater problem. □

André Wyss is in the Department of Geological Sciences, University of California, Santa Barbara, California 93106, USA.

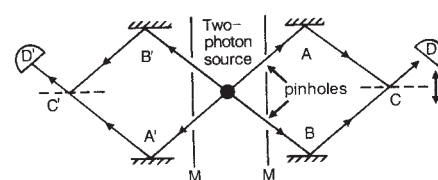


FIG. 1 Two-particle interferometer used by Rarity and Tapster. Photon pairs are radiated from an extended down-conversion source, and reach the detectors (D, D') via pinholes (Young's slits) in the optical masks (M) and the mirrors (A, B and A', B') and half-silvered mirrors (C, C'). When C (say) is translated uniformly, the rate of coincident arrivals at the two detectors oscillates sinusoidally, even though the arrival rate at each detector separately is unaffected.

Einstein. The inequality and hence these commonsense ideas were previously refuted by a long series of experiments on the polarizations of spatially separated particles. (Two^{9,10} of the polarization experiments also used pairs of photons produced by down-conversion.) The new two-particle interferometer experiment of J. Rarity and P. Tapster¹ is the first refutation of Bell's inequality not using polarization. Finally, there exists the possibility of further generalization to interferometry with three, four or more particles¹¹. This would make possible an even more emphatic refutation of Einstein's ideas than provided by Bell's inequality^{11,12}.

Figure 1 (modified for simplicity) shows the arrangement set up by Rarity and Tapster. Photon pairs are emitted at a steady rate by the large-volume down-conversion source and some of these pairs arrive in coincidence at the detectors D and D'. Rarity and Tapster monitored the rate of these coincident arrivals while varying the length of path A from the source to the detector D. The length of path A could be shortened by moving the adjustable mirror up or lengthened by moving it down, and this motion does not affect the lengths of the other paths (B, A', B') to the detectors. They observed that, as the length of path A was varied uniformly in time, the rate of coincident arrivals at the detectors oscillated sinusoidally. This is the essential phenomenon of two-particle interferometry. Furthermore, this sinusoidal oscillation is precisely what Bell showed could not occur if Einstein's ideas about locality and reality were correct.

It is worth emphasizing the quantum mechanics behind this phenomenon. Ordinary single-particle interference occurs whenever a particle is offered two paths from source to detector and the difference in length of these paths is well defined — that is, within a fraction of the de Broglie wavelength of the particle. The two routes are then said to interfere quantum mechanically. For example, each half of Rarity and Tapster's arrangement is, essentially, a Young's double-slit arrangement. Thus, if the source were small, the

1. Gingerich, P. D., Wells, N. A., Russell, D. E. & Shah, S. M. I. *Science* **220**, 403–406 (1983).
2. Gingerich, P. D., Smith, B. H. & Simons, E. L. *Science* **249**, 154–156 (1990).
3. Flower, W. H. *Proc. zool. Soc. Lond.* **1883**, 466–513 (1883).
4. Van Valen, L. *Bull. Am. Mus. nat. Hist.* **132** (1), 1–126 (1966).
5. Szalay, F. S. *Am. Mus. Novit.* **2361**, 1–26 (1969).
6. Scott, W. B. *A History of Land Mammals in the Western Hemisphere* (Macmillan, New York, 1937).

7. Goodman, M., Czelusniak, J. & Beeber, J. E. *Cladistics* **1**, 171–185 (1985).
8. Berta, A., Ray, C. E. & Wyss, A. R. *Science* **244**, 60–62 (1989).
9. Barnes, L. G. & Mitchell, E. D. in *Evolution of African Mammals* (ed. Maglio, V. J.) 582–602 (Harvard University Press, Cambridge, 1978).
10. Fordyce, R. E. *Paleogeog. paleoclimatol. Palaeoecol.* **31**, 319–336.
11. Gould, S. J. & Lewontin, R. L. *Proc. R. Soc. B* **205**, 581–598 (1979).

QUANTUM OPTICS

Two-particle interferometry

Michael Horne, Abner Shimony and Anton Zeilinger

SEVERAL new experiments^{1–3} have demonstrated that the most elementary apparatus for exhibiting quantum behaviour, the single-particle interferometers (related to Young's famous slits and Michelson's mirrors), can be expanded into a remarkable new class of devices: two-particle interferometers. Just as single-particle interferometry demonstrated a fundamental aspect of quantum mechanics — that particles must be described as travelling waves that can be split and remerged — so the new experiments emphasize the fundamental importance of the 'entanglement' of separated particles.

These new interferometers and the phenomena they reveal are exciting for several reasons. First, ordinary single-particle interference is, as Feynman reminded us⁴, "a phenomenon which is impossible, *absolutely* impossible, to explain in any classical way, and which has in it the heart of quantum mechanics". After almost a century of quantum

mechanics, it is surprising to uncover new phenomena so close to its heart^{5,6}. Second, the radiation used in the new experiments, which consists of two photons produced simultaneously from a single photon by a nonlinear process known as down-conversion, is a fascinating quantum mechanical system. Down-conversion radiation, which had been studied intensively both theoretically and experimentally by L. Mandel's group (the authors of ref. 3), was used to produce the first two-particle interferometer effect in a pioneering experiment⁷ in 1987. Some important conceptual questions about the character of down-conversion radiation remain to be investigated.

Furthermore, because the new experiments involve two spatially separated particles behaving in an emphatically quantum mechanical way, they can provide new refutations of Bell's inequality⁸, which is a consequence of certain commonsense ideas about reality and locality (no action-at-a-distance) espoused by

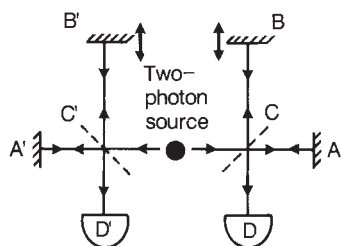


FIG. 2 Two-particle interferometer used by Kwiat *et al.* and Ou *et al.*. The arrangement resembles a pair of Michelson interferometers. The photon going left from the source can reach the detector D' via either the short route off mirror A' or the long route off moveable mirror B'. Similarly for the right-bound photon. The broken lines C and C' indicate half-silvered mirrors. As the adjustable mirrors are moved, the rate of coincident photon arrivals varies, although the signal at the individual detectors is constant.

rate at which single photons arrive at detector D would depend on the relative length of paths A and B — ordinary single-particle interference would occur. But, because the source was large, the relative length of paths A and B was not well defined, and hence ordinary single-particle interference did not occur. Consequently, the rate of arrival at each detector in their experiment was constant independent of the mirror position. Nevertheless, the coincident arrival rate at the two detectors varied with mirror position.

Quantum mechanics accounts for this curious behaviour as follows. Because the two photons are emitted in opposite directions by the source, there are two ways for the two particles to reach the two detectors: the right-going particle takes path A while the left-going particle takes path A'; or the right-going particle takes path B while the left-going takes path B'. In this two-particle interferometer, it is interference between these two possible ways of reaching the detectors that produces the observed oscillations in the coincident arrival rate as the mirror is moved.

A different arrangement was proposed by J. Franson⁵ and used by P. Kwiat *et al.*² and by Z. Ou *et al.*³ (Fig. 2). The arrangement is, again, illuminated by a down-conversion source; and, again, as the mirrors are moved, the coincident arrival rate oscillates sinusoidally, even though the single-photon arrival rate at each detector is constant. For this arrangement, the photon reaching detector D may take the short path via mirror A or the long path via mirror B and similarly for the other photon. But because the two photons are emitted simultaneously and detected simultaneously, it is not possible that one took the long path while the other took the short path. That is, they either both took the long path or both took the short path. The observed oscillations in the coincident arrival rate are caused by quantum mechanical interference

between these two possible ways of reaching the detectors.

In what sense is the phenomenon of two-particle interference close to the heart of quantum mechanics? The answer is that quantum mechanics — but not classical physics — permits a pair of particles to be in an 'entangled state': that is, neither particle by itself is in a definite state, but there are correlations which force the pair to be in a definite, but peculiarly, quantum-mechanical state. As Schrödinger said¹³, entanglement is "not . . . one but rather the characteristic trait of quantum mechanics". For example, entanglement between apparatus and system is the essential feature behind the much discussed problem of measurement in quantum mechanics. Moreover, entanglement of spatially separated particles is the essential feature of quantum mechanics that cannot be accommodated in Einstein's 'local realist' world view.

Can the essence of entanglement be captured in everyday terms? Yes, it can. In Tapster's and Rarity's experiment, each pair of coincident photons observed at the detectors got there by travelling both paths A-A' and paths B-B'. In the Franson-type experiments, each pair of observed coincident photons were emitted both early (taking the long route) and late (taking the short route). The lesson that Bell and the experimentalists have taught us about Einstein's strictly correlated but spatially separated particles is this: any attempt to follow commonsense by substituting 'either-or' for 'both-and' will necessarily fail to match the data from the two-particle interferometer. □

Michael Horne is in the Department of Physics, Stonehill College, North Easton, Massachusetts 02356, USA. Abner Shimony is in the Departments of Philosophy and Physics, Boston University, Boston, Massachusetts 02215, USA. And Anton Zeilinger is in the Institut für Experimentalphysik, Universität Innsbruck, Technikerstrasse 25, A-6020 Innsbruck, Austria.

1. Rarity, J.G. & Tapster, P.R. *Phys. Rev. Lett.* **64**, 2495–2498 (1990).
2. Kwiat, P.G., Vreka, W.A., Hong, C.K., Nathel, H. & Chiao, R.Y. *Phys. Rev. A* **41**, 2910–2913 (1990).
3. Ou, Z.Y., Zou, X.Y., Wang, L.J. & Mandel, L. *Phys. Rev. Lett.* **65**, 321–324 (1990).
4. Feynman, R.P. *Lectures on Physics* Vol. 3 (Addison Wesley, Reading, Massachusetts, 1965).
5. Franson, J.D. *Phys. Rev. Lett.* **62**, 2205–2208 (1989).
6. Horne, M.A., Shimony, A. & Zeilinger, A. *Phys. Rev. Lett.* **62**, 2209–2212 (1989).
7. Ghosh, R. & Mandel, L. *Phys. Rev. Lett.* **59**, 1903–1905 (1987).
8. Ballentine, L.E. (ed.) *Foundations of Quantum Mechanics Since the Bell Inequalities* (Am. Ass. Phys. Teachers, College Park, Maryland, 1988).
9. Shih, Y.H. & Alley, C.O. *Phys. Rev. Lett.* **61**, 2921–2924 (1988).
10. Ou, Z.Y. & Mandel, L. *Phys. Rev. Lett.* **61**, 50–53 (1988).
11. Greenberger, D.M., Horne, M.A., Shimony, A. & Zeilinger, A. *Am. J. Phys.* (in the press).
12. Mermin, N.D. *Am. J. Phys.* **58**, 731–734 (1990).
13. Schrödinger, E. *Proc. Camb. phil. Soc.* **31**, 555–563 (1935).

Feather-light alloy

ICE cubes formed in the refrigerator always have air bubbles in them. The water contains dissolved air, which is expelled on freezing. Daedalus suspects that the same sort of thing happens with metals. The upper part of a casting often has to be rejected as porous; and molten silver dissolves so much oxygen that the metal 'spits' it out as it solidifies.

So he is forcing various gases into molten metals under very high pressure. The pressure is increased still further as the metal solidifies, to keep the gas in solid solution, or at least as very tiny bubbles. The resulting supersaturated metal should be heavily stressed by its loading of occluded gas. On reheating it close to its melting point but at normal pressure, it should creep and expand, allowing the gas to emerge and foam up within it. The result: foamed metal!

DREADCO's 'Foametal'® will have many uses. The crucial requirement for many metal parts and components is not high strength, but high rigidity. Like foamed polystyrene, Foametal will be far lighter than fully dense material, and weight for weight far more rigid. It should be widely welcomed in aerospace engineering, car body construction and ship building (where its ability to float will be a useful bonus). The more highly foamed grades will be readily crushable by collapse of their bubbles, and should make splendid energy absorbing bumpers, helmets and so on. One very useful trick would be to sell Foametal in its gas-loaded, unannealed form. The user would merely put a slug of it in a hot mould, when it would expand, foam up and fill the mould exactly. This 'internal-pressure diecasting' would be ideal for the mass production of small intricate components.

Daedalus also hopes for wide sales to the building trade, and even to do-it-yourself constructors. Foametal should feel and act much like a light but high-tenacity wood. It will be warm and friendly to the touch, and easy to saw or drill. Its crushability will let it accept nails or screws, and be bent or deformed with a hammer. Products made from it will be usefully proof against rot and fire.

Daedalus is also foaming metals with gases other than air. Mixtures containing ammonia might case-harden the inner surfaces of the bubbles by nitriding, giving a very strong Foametal indeed. Ethylene, forced into a low-melting alloy of aluminium, might polymerize under the high pressure and temperature, and fill the metal with polyethylene bubbles. Such a 'plasticized metal' would be heavily damped internally, and would absorb sound wonderfully. It could recreate our whole current crop of noisy machines in much quieter and more civilized form.

David Jones

Calculating greenhouse budgets

SIR—The global warming potentials (GWPs) proposed by Lashof and Ahuja¹ provide a useful step towards a system for converting sources (and sinks) of various greenhouse gases into common greenhouse units. Others² have also derived GWPs, and the Intergovernmental Panel on Climate Change (IPCC) is soon to report on various related issues. But I would like to caution against premature reliance on a GWP scale as a central component of a greenhouse treaty.

Assessing GWPs requires an accurate numerical accounting of three factors for each gas: (1) direct (instantaneous) greenhouse forcing; (2) indirect effects; and (3) atmospheric residence time. The first factor is relatively undisputed, but the other two are not. Regarding indirect effects, the GWPs for chemically active gases — CO, tropospheric O₃, precursors and CH₄ — depend on a complex mix of those gases, NO_x and other factors related to the chemical environment into which they are released. The magnitude of indirect effects varies spatially and temporally³, suggesting the need for GWPs based at least on several chemically coherent regions (such as urban midlatitude, clean midlatitude, tropical) rather than a single, global GWP.

Accounting for residence times is even more complex as behaviour of atmospheric CO₂ differs fundamentally from the other greenhouse gases. Lashof and Ahuja have approached this using Maier-Reimer and Hasselmann's ocean model⁴, from which they calculate an "effective residence time" of 230 years. However, the response to CO₂ input is sensitive to the temporal pattern of emissions (see Table 2 of ref. 4). Under conditions of continuous and exponential growth, long-term response plays little part in determining oceanic uptake. This has been the case for the past few decades, during which roughly half of the CO₂ emitted into the atmosphere has been absorbed — a situation not adequately described using long residence times as uptake is dominated by short-term processes. Conversely, under conditions of declining emissions there is tighter coupling of emissions to longer-term oceanic processes (mixing below the thermocline). Harvey⁵ has argued that the fraction of CO₂ can even be managed at zero (no net increase in greenhouse forcing from CO₂) if emissions are initially reduced by about half. Although the fully integrated "effective residence time" may remain roughly the same under these different conditions, the response functions are clearly different. Thus, the pattern of emissions also affects the sensitivity of a CO₂-based GWP to the relative importance of short and long-term contributors

to global warming (as reflected in the discount rate and the time interval of integration). Perhaps different scales of GWPs should be established for countries with different emission patterns.

Finally, how can GWPs be applied to the political process? GWPs are at present based on somewhat arbitrary assumptions and are fraught with uncertainty and complexity. Results from models of oceanic uptake of CO₂ vary considerably, and Tans *et al.*⁶ suggest an important (but controversial) role for the terrestrial biota in net CO₂ consumption. These carbon cycle factors significantly affect the assumed "effective residence time" for CO₂ and, therefore, the GWP scale, as the non-CO₂ components of the latter are inversely proportional to the former. For example, using the plausible 120-year residence time adopted by Rodhe² instead of 230 years roughly doubles the GWP of N₂O.

Several nations that have acknowledged the utility of GWPs would also like to sign a greenhouse treaty at the 1992 UN Conference on Environment and Development. Given that political schedule and the uncertainty of the scientific basis of GWPs, insistence on use of this parameter will delay and may misdirect a legal regime to control global warming, especially if that regime is highly quantified. Even small shifts in GWPs will have high financial and political consequences. So if they are used, it must be clear well in advance what rules apply to the process of changing the GWP scale.

Despite the theoretical appeal of GWPs, it is important for a greenhouse regime to be relatively insensitive to inevitable fluctuations in them. This probably means that there will have to be independent control plans for CO₂ and for non-CO₂ gases. Moreover, as widely noted^{1,7-9} the absence of reliable, cost-effective measures to monitor emissions of non-CO₂ greenhouse gases (except chlorofluorocarbons) suggests there may be little gained for the goal of a verifiable multi-gas greenhouse regime, even if there is some political agreement on appropriate GWP values.

DAVID G. VICTOR

*Department of Political Science,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA*

1. Lashof, D.A. & Ahuja, D.R. *Nature* **344**, 529–531 (1990).
2. Rodhe, H. *Science* **248**, 1217–1219 (1990).
3. Thompson, A.M., Huntley, M.A. & Stewart, R.W. *J. geophys. Res.* **90**, 9829–9844 (1990).
4. Maier-Reimer, E. & Hasselmann, K. *Clim. Dyn.* **2**, 63–90 (1987).
5. Harvey, L.D.D. *Clim. Change* **15**, 343–381 (1989).
6. Tans, P.P., Fung, I.Y. & Takahashi, T. *Science* **247**, 1431–1438 (1990).
7. Grubb, M. *The Greenhouse Effect: Negotiating Targets* (Royal Institute of International Affairs, London, 1989).
8. Victor, D.G. *Discussion Paper G-90-06* (Energy and Environmental Policy Center, John F. Kennedy School of Government, Harvard University, Cambridge, 1990).
9. Victor, D.G. *Policy Sciences* (in the press).

Origins of fetal handedness

SIR—The origins of behavioural asymmetries — in particular handedness — are a topic of continued interest. We have now found evidence of a behavioural asymmetry in the fetus — a preference for sucking the thumb of the right hand.

In the human population as a whole some 90 per cent exhibit a bias for the right hand¹. In our study we used real-time ultrasound to observe thumb-sucking by the fetus in 224 women who had normal, singleton, uncomplicated pregnancies. Our results, presented in the table, reveal a clear bias for sucking the thumb of the right hand, only 12 of 224 (5.4 per cent) fetuses favouring the left. The

FETAL THUMB PREFERENCE

	Gestational age (weeks)			
	15–21	28–34	36–term	Total
Right thumb	52	76	84	212
Left thumb	4	3	5	12
z	6.28	8.11	8.27	13.302
P<	0.0001	0.0001	0.0001	0.0001

The number of individuals observed sucking either their left or right thumb at different gestational ages ($n = 224$). z, Right thumb bias.

intrauterine position of the fetus had no effect on thumb preference. Repeated observations of 17 fetuses indicated that preference for a particular thumb is maintained during pregnancy. Furthermore, fetal thumb preference correlates highly with neonatal head-position preference ($P = 0.002$).

Our study demonstrates for the first time the existence of behavioural asymmetries before birth. Evidence of handedness in fetuses of normal uncomplicated pregnancies raises questions for theories proposing that laterality results from prenatal or birth injury² but supports genetic explanations of handedness³. One could speculate that although the initial bias is a result of advanced right-sided neuromuscular development, its maintenance results from reinforcing aspects of thumb-sucking. Indeed it may be that sucking a particular thumb differentially stimulates the left and right brain and thus commences the development of brain lateralization prenatally.

PETER G. HEPPER
SARA SHAHIDULLAH
RAYMOND WHITE

*Fetal Behaviour Investigations,
School of Psychology,
Queen's University of Belfast,
Belfast BT7 1NN, UK*

1. Porac, C. & Coren, S. *Lateral Preferences and Human Behavior* (Springer, New York, 1981).
2. Satz, P. *Neuropsychologia* **11**, 115–117 (1973).
3. Annett, M. *Left, Right, Hand and Brain: The Right Shift Theory* (LEA, Hillsdale, New Jersey, 1985).

Ocean currents and climate

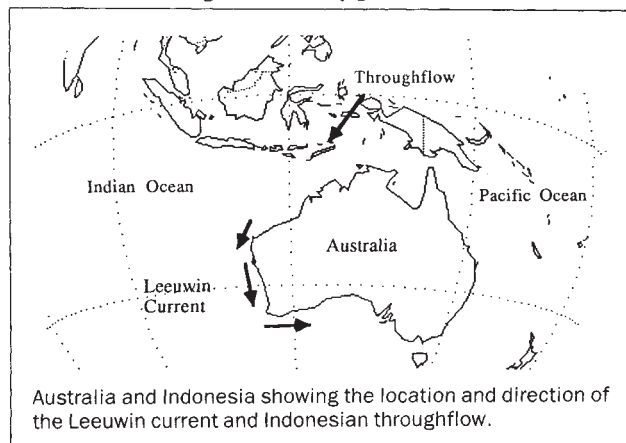
SIR—The Leeuwin current, an anomalous eastern boundary current, flows poleward along the west coast of Australia into the prevailing southerly winds. Its existence relies on the presence of channels through the Indonesian Archipelago, allowing warm western equatorial Pacific waters to flow through to the north-west tip of Australia (see figure). It is unlike the other eastern boundary currents of the world oceans (for example, California, Peru, Canary, Benguela) in that there is neither surface equatorward flow nor coastal upwelling¹. These other eastern boundary currents are all part of a large-scale anticyclonic gyre driven by anticyclonic winds. Here I wish to draw an analogy between the climate effects of the Leeuwin current and the effect of the closing of the isthmus of Panama 3.2 million years ago.

The Leeuwin current is driven by a deep alongshore density gradient in the Indian Ocean²⁻⁵. This alongshore density gradient

current^{6,8}. At other eastern boundaries, the equatorial surface temperatures are much cooler and convective overturning does not occur until polewards of about 40° or 50° N or S and so the circulation is dominated by local wind forcing and the flow is equatorwards.

The Leeuwin current has profound effects on the climate of western Australia. Very warm waters are advected polewards by the current and there is a large heat loss to the atmosphere poleward of 20° S where convective cooling occurs⁶. Perth and the rest of southwestern Australia experience a very mild and wet climate, receiving up to 50 inches of rain a year in places. This is in contrast to climates at similar latitudes on other eastern continental boundaries, for example, Baja peninsula, central Chile and Morocco, which are all very arid with less than 10 inches of rain a year.

Before about 3.2 million years ago, there was a channel through the isthmus



drives a strong onshore geostrophic flow which turns and flows polewards on hitting the coast of western Australia. A passage through the Indonesian Archipelago is crucial for the existence of the deep alongshore density gradient and hence the Leeuwin current⁶. Warm, western equatorial Pacific waters flow through the archipelago and maintain very high steric heights off northwestern Australia. Internal Kelvin waves propagate along the coast of western Australia, radiating internal Rossby waves westward, thus tending to maintain specific volume anomaly profiles off the coast of western Australia equal to those at the equatorial end of the boundary. Polewards of about 20° S, convective cooling⁷ occurs and the deep alongshore density gradient is set up. In a sense, therefore, the Leeuwin current can be thought of as being driven by equatorial Pacific winds which pile up warm water in the western equatorial Pacific. Furthermore, the actual magnitude of throughflow from the Pacific to Indian oceans, through the Indonesian Archipelago, has little influence on the Leeuwin

of Panama connecting the Atlantic and Pacific oceans, much like the channel connecting the Indian and Pacific oceans observed today (Fig. 1). By analogy with the Leeuwin current, I suggest that an intense poleward flowing reverse California current existed which accelerated into the opposing winds. This current would have had a profound effect on the marine climate and environment all along the coast of California up to higher latitudes. In particular, a strong northward heat transport would occur along the western coast of North America, and a large heat loss to the atmosphere would occur poleward of about 20° N. Along Baja and California, the climate would have been milder and wetter, similar to the climate of Perth today. In the coastal ocean, upwelling would be suppressed because the current would flow into the prevailing winds, as observed off western Australia. Thus, the nutrient abundance in the surface waters would probably be depleted, with the result that there would be a significant drop in phytoplankton and primary production.

With the closure of the isthmus of Panama, the situation would become comparable to what is observed today. The current would reverse and flow in the direction of the prevailing northerly wind. Intense upwelling would begin along the coast of North America and there would be a significant drop in the poleward heat transport in the Pacific Ocean. I also propose that a similar phenomenon may have

occurred along the coast of South America where the equatorward-flowing Peru Current exists today.

ANDREW J. WEAVER

Department of Meteorology
McGill University,
Montreal, Quebec H3A 2K6,
Canada

1. Wooster, W.C. & Reid, J.L. in *The Sea* Vol. 2, 253–280 (Wiley, New York, 1989).
2. McCreary, J.P., Shetye, S.R. & Kundu, P.K. *J. mar. Res.* **44**, 71–92 (1986).
3. Thompson, R.O.R.Y. *J. mar. Res.* **45**, 7813–7827 (1987).
4. Weaver, A.J. & Middleton, J.H. *J. phys. Oceanogr.* **19**, 626–648 (1989).
5. Weaver, A.J. & Middleton, J.H. *Cont. Shelf Res.* **10**, 105–122 (1990).
6. Godfrey, J.S. & Weaver, A.J. *Prog. Oceanogr.* (in the press).
7. Hanley, R.L. *J. phys. Oceanogr.* **1**, 241–248 (1971).
8. Kundu, P.K. & McCreary, J.P. *J. phys. Oceanogr.* **16**, 2191–2198 (1986).
9. Crowley, T.J. *Rev. geophys. Space Phys.* **21**, 828–877 (1983).
10. Crowley, T.J. in *Physically-Based Modelling and Simulation of Climate and Climatic Change Part II*, 883–949 (Kluwer, Dordrecht, 1988).

Methane in cities

SIR—The proposed control measures designed to reduce the climatic impact of greenhouse gases have largely been confined to cessation of release of chloro-fluorocarbon gases, and less wasteful use of energy derived from the combustion of fossil fuels. We believe that improved controls on the transmission of natural gas, particularly in eastern Europe and Asia, offer a promising opportunity to slow the global build-up of methane, another significant greenhouse gas.

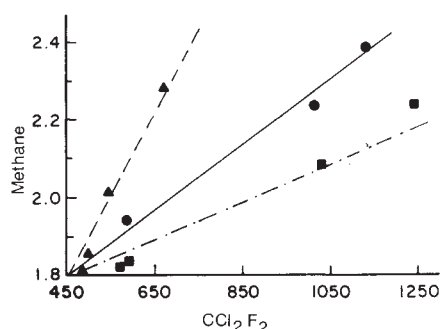
Because the atmospheric lifetime of methane is about 10 years¹, its observed average increase of 1 per cent per year during the 1980s (ref. 2) implies a worldwide excess of only 10 per cent of methane sources over sinks. Therefore, a 10 per cent reduction in the total emissions of methane should stabilize its global concentration and eliminate any further incremental contribution of this gas to the atmospheric trapping of infrared radiation. The increase in methane concentrations has been estimated³ to have contributed about 20 per cent of the incremental infrared greenhouse forcing since 1850.

We collected 27 whole-air samples into evacuated two-litre canisters at ground level, remote from point sources or vehicular traffic, and returned them to our laboratory for gas chromatographic analysis. Separate aliquots were analysed for methane, carbon monoxide, halocarbons and non-methane hydrocarbons from C₂ to C₅. Usually, we included a site 10–25 kilometres upwind of the city centre to estimate the increase in concentration attributable to the urban area itself. Methane in air samples collected in May 1990 from four eastern European cities (Berlin, Budapest, Krakow and Prague) had substantially higher concentrations

than those characteristic of most US cities, or of two other eastern European cities (Vienna and Zagreb). Because most other major global methane sources are biological and non-urban (such as rice, cattle or swamps)⁴, leakage during natural-gas transmission is the most probable source of this excess urban methane. Furthermore, air samples collected from the Yerevan region of Armenia in January 1990 exhibited even higher concentrations of methane (3.3–3.5 parts per million per volume) and most other light hydrocarbon gases (0.02–0.1).

Although the excess methane concentrations in the European cities were clearly of local urban origin, the hydrocarbon concentrations 15 kilometres away from Yerevan were as large as those in the city itself, suggesting regional release over a broader area. At that time, industrial activity in Armenia was only about one-quarter of normal because of disruptions for transport.

Most chlorofluorocarbon gases are emitted directly into city air and are excellent markers of urban contamination



Correlations of measured urban concentrations of methane in parts per 10^6 with CCl_2F_2 in parts per 10^{12} in three cities. Squares, Budapest; circles, Prague; triangles, Krakow.

of an air mass⁵. The increase in methane concentrations versus those of CCl_2F_2 are shown in the figure for Budapest, Krakow and Prague. In each city, the incremental increase in methane is approximately proportional to that for CCl_2F_2 , demonstrating the local urban origin for methane. The intercepts in the figure are consistent with the concentrations found in remote regions for these latitudes and season.

F. SHERWOOD ROWLAND
NEIL R. P. HARRIS
DONALD R. BLAKE

Department of Chemistry,
University of California,
Irvine,
California 92717, USA

1. Mayer, W.E. et al. *Proc. natn. Acad. Sci. USA* **79**, 1366–1370 (1982).
2. Blake, D.R. & Rowland, F.S. *Science* **239**, 1129–1131 (1988).
3. Ramanathan, V., Cicerone, R.J., Singh, H.B. & Kiehl, J.T. *J. geophys. Res.* **90**, 5547–5566 (1985).
4. Cicerone, R.J. & Oremland, R.S. *Biogeochem. Cycles* **2**, 299–327 (1988).
5. Blake, D.R., Woo, V.H., Tyler, S.C. & Rowland, F.S. *Geophys. res. Lett.* **11**, 1211–1214 (1985).

Dynamical mass-transfer paradox

SIR—Although the periodically varying frequency of the emission lines of the star SS433 can be described kinematically¹ by two precessing jet sources emanating from one component of a binary system, many of its physical and dynamical properties remain enigmatic. In particular, the jet speed and power require a compact source whereas the precessional behaviour would be appropriate for a slowly rotating giant star. Here we summarize this dichotomy and speculate on a novel resolution of it.

The terminal speed v_∞ of stellar mass loss is characteristically of the order of the escape speed, so the value $v_\infty = 0.26 c$ implies that SS433 is a neutron star or black hole. Further, only gravitational accretion onto a collapsed object suffices to power the jets ($10^{30} \text{ erg s}^{-1}$). The thick supercritical keplerian disk required for accretion from the companion has been proposed² as providing jet collimation and invoked to explain the observed³ occultation of the X-ray jets within 10^{12} cm .

A keplerian disk, however, cannot as a whole undergo coherent driven precession, and it has been proposed⁴ that the disk geometry is somehow 'slaved' by angular momentum of matter incident from the giant companion, which can precess coherently. This poses two problems. First, for the companion to provide the observed amplitude and phasing of the 'slave' precession pattern, it must be a slow rotator^{5,6}. Proper analysis of this suggestion must include the orbital as well as the rotational angular momentum. Second, angular momentum transfer in the accretion stream takes at least the free-fall time (about 6 days) which is comparable⁷ to the jet nodding period. Therefore, the jet axes move as if they emerged from the slow rotation axis of a giant star undergoing driven precession⁸, though their speed and power require they originate from a collapsed object.

To reconcile this dichotomy, we conjecture that the jets emanate from a polar hole in a precessing giant star in whose core resides an accreting collapsed object which powers them. Such a hybrid object could form by rapid dynamical mass transfer from a giant before helium ignition⁹ onto an already critically accreting collapsed object, the polar hole being blown by a pre-existing jet and sustained in the rotating giant star as described in ref. 9. We suggest first that mass expulsion from the central accretor will occur preferentially along the reduced overpressure direction of the hole and, second, that the hole will follow the driven precession of the hybrid. As regards the second suggestion, enough time has elapsed since the formation of SS433 for circulation to lead to uniform rotation of the hybrid and

hence allow its entire structure to precess coherently^{10,11}. In our picture, occultation of the X-ray jets³ is caused by the hybrid itself (with a radius of about 10^{12} cm) rather than by a thick disk.

We believe our preliminary model has possible attractions beyond reconciling the jet speed and precession properties. For example, the stability of the radiation field in the funnel may explain the remarkable jet speed constancy. The model should also be amenable to predictive observational tests, including the eclipse effects of two large stars.

G. W. COLLINS II

Department of Astronomy,
Ohio State University,
Columbus, Ohio 43210, USA

J. C. BROWN

Department of Physics and Astronomy,
University of Glasgow,
Glasgow G12 8QW, UK

J. P. CASSINELLI

Department of Astronomy,
University of Wisconsin,
Madison, Wisconsin 53706, USA

1. Abell, G.O. & Margon, B. *Nature* **279**, 701 (1979).
2. Abramowicz, A. & Piran, T. *Astrophys. J.* **241**, 67 (1980).
3. Stewart, G.C., Pan, H.-C. & Kawai, N. *Proc. 23rd ESLAB Symp. ESASP-296* (1989).
4. van den Heuvel, E.P.J., Ostriker, J.P. & Petterson, J.A. *Astron. Astrophys.* **81**, L7 (1980).
5. Collins, G.W. II & Newsom, G.H. *Astrophys. J.* **308**, 144 (1986).
6. Collins, G.W. II & Newsom, G.H. *Astrophys. J.* **331**, 486 (1988).
7. Begelman, M. & Rees, M.J. *Mon. Not. R. astr. Soc.* **206**, 209 (1984).
8. Webbink, R.F. *IAU Symp.* **73**, 35 (1985).
9. Amendt, P., Lanza, A. & Abramowicz, M.A. *Astrophys. J.* **363**, 437 (1989).
10. Mestel, L. & Takhar, H.S. *Mon. Not. R. astr. Soc.* **156**, 419 (1972).
11. Papaloizou, J.C.B. & Pringle, J. *Mon. Not. R. astr. Soc.* **201**, 49 (1982).

Pebble shape

SIR—Based on his measurements of the proportions of the three principal diameters of beach pebbles, Wald in Scientific Correspondence¹ concluded that the shapes of pebbles evolve toward the stable form of a flattened ovoid due to abrasion by waves. In response, Ashcroft² attributed the flat shapes of beach pebbles to planar structures in the original rock material.

This fascination with pebble shapes has existed for more than 150 years. In 1834, Palmer³ discussed the role of waves in sorting pebbles on the beach according to their sizes and shapes. Writing in 1898, Cornish⁴ suggested the shapes and sizes of beach gravels are initially a product of the natural fragmentation of the original rock material, but abrasion by waves then modifies their shapes. In a paper published by *Nature* in 1944, Lord Rayleigh⁵ reported his laboratory experiments simulating the abrasion of pebbles lacking internal planar structures. In those experiments, he could reproduce flattened disks very similar to those found on beaches.

More recent investigators have used

other approaches.

One enlightening study was that by Dobkins and Folk⁶ of pebbles on beaches in Tahiti. Among the advantages of studying beaches in Tahiti is that all the pebbles are of essentially one rock type, basalt, which has no bedding or crystal orientation and so tends to abrade uniformly. Dobkins and Folk found that the pebbles on the beaches are systematically flatter than those in the streams that originally carried them to the shore, and suggested that this difference results predominantly from pebbles' sliding motion back and forth in the wave swash and from sand blasting. On beaches composed of mixtures of pebbles and sand, slightly flattened pebbles tend to lie on the beachface with their sides upward. The sand carried by the wave swash mainly hits the pebbles on their exposed sides, chipping out minute pits and slowly increasing the degree of flatness. At the same time, spherical pebbles are easily rolled down the beach slope by the backwash of waves. This tends to carry the spherical pebbles offshore, leaving the flatter pebbles on the upper beach where they are more likely to be observed.

Clearly, fragmentation along planar structures in the original rock material can yield flattened pebbles. Equally apparent is that wave-induced abrasion is important in producing the marked symmetry and flat shapes of beach pebbles.

MARK S. LORANG

PAUL D. KOMAR

College of Oceanography,
Oregon State University,
Corvallis,
Oregon 97331, USA

1. Wald, Q. R. *Nature* **345**, 211 (1990).
2. Ashcroft, W. *Nature* **346**, 227 (1990).
3. Palmer, H. R. *Phil. Trans. R. Soc.* **124**, 567–576 (1834).
4. Cornish, V. *Geogr. J.* **11**, 628–651 (1898).
5. Rayleigh, Lord *Nature* **154**, 169–171 (1944).
6. Dobkins, J. A. & Folk, R. L. *Sediment. Petrol.* **40**, 1167–1203 (1970).

Mosaic resistance in plants

SIR—The existence of genetic variation between different parts of individual plants has long been exploited by plant breeders. This variation could be the basis for an adaptive response by plants to environmental change via the proliferation of tissues expressing newly arisen mutations. Such a process would lead to changes in the phenotype of individual trees over time as well as to the production of seed from genetically different branches. We have obtained the first direct evidence of an adaptive role for genetic variation within plants by measuring the composition of volatile oils among branches of eucalypt trees, the variation of which confers differential resistance to a defoliating insect.

Christmas beetles inflict heavy damage on trees of the genus *Eucalyptus* in south-east Australia, particularly in areas disturbed by agriculture (see ref. 1 for a review). Tree clearance and increased areas of improved pastures provide ideal habitat for soil-dwelling Christmas beetle larvae and this has resulted in a high level of herbivore pressure on the remaining trees when the adults emerge from the soil. During a beetle outbreak near Yeoval, New South Wales, we found a specimen of yellow box (*Eucalyptus melliodora*) which had a conspicuous branch immune to attack by a species of Christmas beetle (*Anoplognathus montanus*). By contrast, other branches were infested with beetles which completely defoliated the rest of the tree (see figure). This observation was supported by feeding trials in the laboratory where beetles were fed exclusively with leaves from resistant or susceptible branches.

A comparison of the volatile oil content of leaves from resistant and susceptible branches using gas chromatography showed major differences in the proportions of the main oil components, and these were consistent within each foliage type. The generality of the phenomenon observed in this single specimen of *E. melliodora* was supported by an examination of other trees of the same species in the same area. Of 20 nearby trees, two were uniformly susceptible to attack by *A. montanus* whereas the rest were uniformly resistant. The biochemical differences between these trees were similar to the differences between the susceptible and resistant branches of the mosaic tree.

We found no other mosaic *E. melliodora* at Yeoval, but examples have been discovered at other sites, and again similar differences in oil composition between resistant and susceptible branches were recorded. Furthermore, examples of within-tree variability involving similar biochemical differences have now been found in three other species of *Eucalyptus*.

The pattern of within-tree variation observed in *E. melliodora*, involving entire branches, strongly suggests that the resistant plant tissue has developed from meristematic cells containing newly arisen somatic mutations. Indeed, the biosynthesis of volatile oils is known to be under genetic control (see, for example, refs 2,3), although environmental and developmental factors can also influence the expression of various components (for example refs 3,4).

The occurrence of mosaic *E. melliodora*



Most of this *E. melliodora* tree has been completely defoliated by *A. montanus*, but one branch (on the left) remains virtually untouched.

may be a specific response in susceptible trees to the increase in insect activity brought about by habitat disturbance, and could thus provide a mechanism for long-lived species such as trees to persist in the face of herbivore pressure. In addition, resistant branches will produce seed-containing genes for resistance which will be subject to selection among seedlings. We are raising seedlings from seed obtained from the resistant and susceptible branches to examine their oil content.

PENELOPE B. EDWARDS

W. J. WANJURA

W. V. BROWN

Division of Entomology,
CSIRO, GPO Box 1700,
Canberra, ACT 2601, Australia

JOHN M. DEARN

Applied Ecology Research Group,
Faculty of Applied Science,
University of Canberra,
PO Box 1, Belconnen, ACT 2616,
Australia

1. Ohmart, C. P. & Edwards, P. B. A. *Rev. Ent.* (in the press).
2. Hanover, J. W. *Heredity* **21**, 73–84 (1966).
3. Lincoln, D. E., Murray, M. J. & Lawrence, B. M. *Phytochemistry* **25**, 1857–1863 (1986).
4. Mihaliak, C. A., Couvet, D. & Lincoln, D. E. *Biochem. System. Ecol.* **17**, 529–533 (1989).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references. □

Unscientific postscript

John C. Marshall

Ludwig Wittgenstein. A Student's Memoir. By Theodore Redpath. Duckworth: 1990. Pp.109. £12.95.

SOME forty years after his death, the cult of Ludwig Wittgenstein burns more fiercely than ever. Scarce a week goes by but that some elderly gentleman remembers (and publishes!) a chance remark that Wittgenstein made when tying his shoelaces; for every (relatively lucid) sentence that Wittgenstein wrote there must now be thousands of pages of (exceedingly dense) exegesis in print. Here, then, is one of the profound mysteries of the twentieth century: how did a minor Viennese aphorist come to be regarded (in some circles) as a great philosopher who had *twice* changed the course of that discipline, first with the *Tractatus* (1921) and then the *Philosophical Investigations* (1953)?

Theodore Redpath attended Wittgenstein's Cambridge 'happenings' (they can hardly be called lectures or seminars) from 1934 to 1936, and from 1938 to 1940. By then, the myth of the shaman was already well-established. I. A. Richards recommended attendance at the shrine in Trinity College with the claim "that Wittgenstein lived like a saint, that he was 'beautiful', but that he wore himself out during the course of a term's lecturing and that the ravages the effort made on him were terrible". And all this, Redpath reports Wittgenstein as saying at the first meeting, merely to help cure any students who might be "suffering from a particular kind of mental cramp".

Redpath himself seems to have been taken in rather less than most; to him, Wittgenstein looked "more like an active general than the mystical philosopher whose appearance I had concocted in the crannies of my brain". But Redpath seems also to have missed the point of the performance: for this is the Wittgenstein who tried to convince his youthful disciples

that it was an error of grammar to imagine that appearances could be represented in the brain (or the mind). To this very day, the acolytes of north Oxford preach that current cognitive psychology, artificial intelligence and neuroscience are all based upon an awful mistake: the



Ludwig Wittgenstein — has twice changed the course of philosophy.

purported failure to realize that language games, not empirical (and theoretical) inquiry, explicate the mental.

When Wittgenstein was still a young engineer, a brief chat with the likes of Alan Turing or Kenneth Craik could (perhaps) have saved him. Sadly, by the time the three men were simultaneously resident in Cambridge, Wittgenstein was living his own legend. His mental cramp had ossified to such an extent he could not see that his approach to psychology was strictly analogous to arguing that astronomy should be bounded by analysis

of the sentence 'The Sun rises in the east and sets in the west'.

How this tragic waste of a fine intellect came about would form a fascinating biography. Yet the first volume of the official biography (by Brian McGuinness, published by Duckworth in 1988) sheds little light on the early development of Wittgenstein's personality, and more by what is suppressed than by what is included.

When Wittgenstein left Vienna for Manchester and Cambridge before the 1914–18 war, to what extent was he avoiding competition with the intellectual and artistic genius of the Hapsburg capital?

Was he later aware of recreating at Trinity the kind of superheated circus that once gathered around Arthur Schnitzler at the Café Griensteidl, or Peter Altenberg at the Café Central? Did he intentionally choose students in a strange land who would be incapable of understanding who he was and what he taught? Redpath recounts how Wittgenstein (normally tieless) borrowed tails from him for a formal dinner at Trinity. Afterwards, "he had the laundering done by the Wellbrook Laundry at Girton. Wittgenstein thought highly of it, and later my family transferred to it." That it should come to this for a man in whose family house, on the Allee-gasse, Johannes Brahms and Clara Schumann had played piano duets.

But the Empire did not pass away without pain. As an artillery officer on the eastern front, what guilt did Wittgenstein incur as he observed, indeed participated in, the slaughter of his people when the armies of the Emperor and the Tzar swept over the defenceless shtetlach of Galicia and Bukovina? Little wonder, then, that the *Tractatus*, drafted in field-pencil at the front, should end: "*Wovon man nicht sprechen kann, darüber muss man schweigen*" (Whereof one cannot speak, thereof one must be silent). □

John C. Marshall is at the Neuropsychology Unit, University Department of Clinical Neurology, The Radcliffe Infirmary, Oxford OX2 6HE, UK.

Footsteps to a solution

Martin Lockley

Dinosaur Tracks. By Tony Thulborn. Chapman and Hall: 1990. Pp. 410. £35, \$85.

THE Chinese philosopher Lao Tse's journey of a thousand miles began with a single step. For the rapidly expanding science of dinosaur tracking, or dinosaur ichnology, this initial step was taken in 1989 with the publication of *Dinosaur Tracks and Traces* (Cambridge University Press) edited by myself and D. D. Gillette. This year, Tony Thulborn has taken the second step, helping to coax this maturing discipline in the right direction. *Dinosaur Tracks* is a detailed, primarily literature-based survey of traditional biological interpretations that have emerged from the study of dinosaur footprints over the past century and a half.

One of the strengths of Thulborn's book is that it contains a wealth of information under the same cover. But this is also one of the book's weaknesses, particularly in the chapters on historic discoveries and principal track types. Here the reader is bombarded with repetitive information, much of it couched with disclaimers about how much uncertainty exists in the assignment of tracks to particular dinosaurs.

Essentially, Thulborn has emulated the model of several German ichnologists, such as Haubold and Abel, who have published encyclopaedic catalogues of

fossil footprints. He has departed from this tradition mainly in the inclusion of two crisply written introductory chapters, an in-depth analysis of dinosaur size, gait and speed, some cautionary tales, and a final chapter on assemblages of dinosaur tracks. This mix adds breadth to the book, but not coherence.

Thulborn's original research shows through most clearly in the later chapters



Dinosaur trackways shown in an aerial view of part of the Purgatoire River in Colorado. The tracks indicate a sauropod-theropod fauna. (Reproduced from *Dinosaur Tracks and Traces*, by Martin Lockley and D. D. Gillette, Cambridge University Press.)

where he presents exhaustive morphometric data and analyses of dinosaur size, gait and speed based on a review of trackway and skeletal data. Since he first aired the speed debate in *Nature* (292, 273; 1981), Thulborn has made a substantial contribution in this area, and his modifications and refinements of R. McN. Alexander's popular formula for the estimation of speed (*Nature* 261, 129; 1976) should be followed by researchers whenever appropriate. On this subject it is

surprising that Thulborn does not take Robert Bakker to task for his sloppy and off-the-cuff claims about "dinosaur cruising speed" made in the book *Dinosaurs Past and Present*.

One of the more disappointing aspects of Thulborn's book is that it perpetuates too many long-standing myths and accepts many fanciful and ill-founded interpretations at face value. For example, the author too often makes sweeping generalizations about how it is rare to find tracks and bones together — but this is not necessarily the case. Thulborn also claims that casts are more durable than moulds, and frequently cites tracks as evidence of dinosaur swimming behaviour. This latter notion has been one of the most misleading models ever to emerge from dinosaur ichnology. Similarly, hypotheses about predator-prey skirmishes and structured herds are handled incompletely, with a curious blend of caution and acceptance of published dogma. Thulborn also unfairly characterizes Edward Hitchcock's pioneer work as wrong, and in so doing glosses over the now well-established relationship between dinosaurs and birds.

Although Thulborn has clearly delved deep into the literature, he has also apparently overlooked or avoided various recent books, including the two I have already mentioned. It is disconcerting to find a source referred to in one context but ignored in another. In some cases, these inconsistencies have resulted in incorrect identification of tracks at particular sites.

Thulborn does not deal much with topics of geological or sedimentological significance. His comments on footprint preservation generally make sense, though he does not say much about underprints or ghost prints. The stratigraphical and geographical location of footprints is also, despite its importance, covered with astonishing brevity, but as Thulborn warns, "most of the book is concerned with footprints as sources of information about dinosaurs", and not about the application of footprints in such areas as biostratigraphy, facies analysis or the problems of preservation.

In fairness, Thulborn has tackled a difficult and frequently neglected discipline that has too often in the past suffered from lack of scientific rigour. His estimations of dinosaurs' size, gait and speed from footprints will provide trackers with a useful repertoire of guidelines to apply and test, and for this reason alone specialists will want a copy of this book. Most of the other areas that Thulborn touches on remain in need of thorough analysis, a task that should not be difficult considering the landslide of new information now emerging.

The recent resurgence of interest in dinosaur ichnology has been fuelled by the discovery of hundreds of new sites throughout the world. This establishment

New in paperback

■ *Geomorphological Techniques* has now been issued in a second edition. The volume is edited by Andrew Goudie and six collaborators for the British Geomorphological Research Group and provides a manual of useful techniques for geomorphologists and others in the Earth and environmental sciences. Publisher is Unwin Hyman, price £24.95.

■ *BLymphocytes* by G. G. B. Klaus is published by IRL Press at Oxford University Press as part of the 'In focus' series. Written to keep students up to date with this fast-moving field, price is £6.50.

■ New from Gordon and Breach is *The Newtonians and the English Revolution 1689–1720* by Margaret C. Jacob, who describes the integration of Newton's science with the interests of the Anglican church. Price is \$32 (\$19 for members of the Science and Arts Society).

■ *Main Group Chemistry* is an introduction to this diverse topic for students and teachers. By A. G. Massey, the publisher is Ellis Horwood and the price £16.95.

■ In *The Fisherman's Problem*, Arthur F. McEvoy describes his research on the ecological and legal background of California's fisheries from 1850 to 1980. Publisher is Cambridge University Press, price £12.95, \$14.95. □

of a large database has been the first step on our journey towards a new understanding. Compilation of some of this information in *Dinosaur Tracks* and in *Dinosaur Tracks and Traces* is a useful second step. Trackers are now in motion and will begin to hit their stride once this avalanche of information is better synthesized and more rigorously interpreted. □

Martin Lockley is in the Department of Geology, University of Colorado, 1200 Carmer Street, Denver, Colorado 80204, USA.

Orbital view

Nicholas E. White

Imaging X-ray Astronomy: A Decade of Einstein Observatory Achievements. Edited by Martin Elvis. Cambridge University Press: 1990. Pp.348. £30, \$49.50.

THE launch of NASA's Einstein observatory on 13 November 1978 thrust X-ray astronomy into the main stream of astrophysics and caused a revolution in the prevailing view of the Universe. Before then, X-ray astronomy had been the domain of a select band of astrophysicists studying a few hundred X-ray sources using crudely collimated detectors with fields of view of a few degrees. The grazing incidence imaging optics of the Einstein observatory gave for the first time in the X-ray domain an angular resolution of a few arcseconds. This resolution, combined with an impressive array of imaging and spectroscopic detectors, produced a wealth of data still largely unsurpassed ten years later. The dramatic increase in sensitivity resulted in the discovery that virtually every class and type

of astronomical object emits X-rays "from Jupiter in our solar system, to distant quasars near the edge of the observable universe", to quote from this volume, published to celebrate the tenth anniversary of the launch of the Einstein observatory.

The book contains a series of review articles and contributed papers and is based on a symposium held at the Center for Astrophysics, Cambridge, Massachusetts. The book begins with brief personal recollections by some of the astrophysicists involved in the original proposal made to NASA in the late 1960s for a large orbiting X-ray telescope. The accounts provide a glimpse into how the determination and foresight of a few eventually prevailed on NASA to launch Einstein, a scaled-down version of the original proposal, ten years after it was first suggested.

The book is split into three parts. The first begins with an excellent review by Harvey Tananbaum summarizing the main results from the Einstein observatory, and discussing the prospects for NASA's next major X-ray satellite, the advanced X-ray astrophysics facility (AXAF). The rest of this part provides more detailed reviews by distinguished astrophysicists about the Einstein results on stellar coronae, cataclysmic variables, X-ray binaries, supernova remnants, active galactic nuclei, clusters of galaxies, nearby galaxies and the all-sky surveys. Most of these provide a full, well-written overview of the state of their respective fields in 1989. Notably absent, however, are detailed reviews of the Japanese and European X-ray astronomy missions that have followed and built on the initial Einstein discoveries.

Part 2 contains contributed papers

AIDS TARGETED INFORMATION/ATIN

TARGETS YOUR LITERATURE SEARCH

Expand your base of knowledge about AIDS/HIV with *AIDS Targeted Information/ATIN*. Every month, *ATIN* not only provides abstracts but also in-depth evaluations of the current published scientific literature on AIDS. Written by clinicians and researchers for clinicians and researchers, *ATIN* provides an authoritative command of the world's literature on AIDS. Published by Williams & Wilkins. Indexed. 12 issues per year. \$125.

To order call
1-800-638-6423

Sponsored by the American Foundation for AIDS Research (AmFAR).

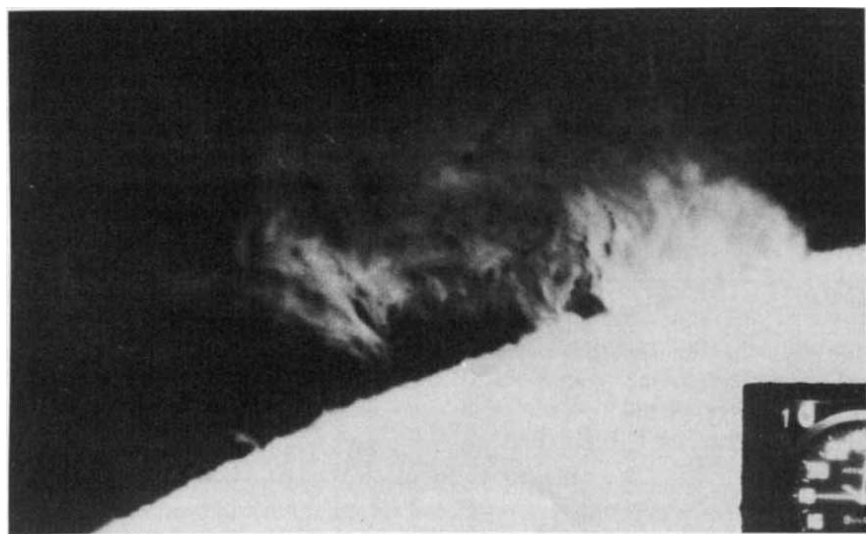
Reader Service No.430

selected to demonstrate the continuing research activities being generated by the Einstein databank. There are a few outstanding papers here ("The Softest AGN" by Cordova *et al.* is one example) but many are of interest only to a few specialists. Fortunately, these take up a relatively small fraction of the book.

Part 3 describes catalogues (referred to as databases) that are being produced for the various classes of objects observed by Einstein. This part consists mostly of a series of brief summaries of the plans to make the catalogues widely available. It is superficial, and mostly an advertisement of things available elsewhere or yet to come.

In summary, this book is an above-average conference proceedings. The first two-thirds is by far the best and provides a good historical reference to how astrophysics was revolutionized by the launch of the Einstein observatory. But the lack of a balanced view of the X-ray observations that followed in the 1980s circumscribes it as a more general reference to the state of X-ray astronomy in 1989. Nonetheless, it still provides an excellent starting point for those interested in learning more about this important X-ray observatory. □

Nicholas E. White is at the EXOSAT Observatory, ESTEC, 2200 AG Noordwijk, The Netherlands.



Solar prominences, photograph taken from the Big Bear Solar Observatory, California, 26 May 1978 — an example of what the experts can do. In *The Amateur Astronomer* Patrick Moore offers advice to both absolute beginners and those with a little experience. He tells what equipment to use, how to use it and what to look for with it. (New, updated edition, published by Norton, price is \$35.)

Core text

S. A. Morse

Early Precambrian Basic Magmatism.
Edited by R. P. Hall and D. J. Hughes.
Blackie: 1990. Pp.486. £85, \$141.

TERRA est omnis divisa in partes tres: core, mantle and crust. The core is metal and mostly liquid, in which fluid flow generates our magnetic field. The mantle, about 63 per cent of the Earth's mass, is dominantly magnesium-iron (mafic, basic) silicate rock, and the crust (5 per cent) is mostly silicon-aluminium (felsic, acid) silicate rock. Plastic upwelling in the mantle conveys heat from the depths of the Earth, and ultimately drives the tectonic motion of the rigid plates. The decompression of hot mantle rock in these upwellings causes melting to occur at some places and times, and the melt is mafic magma that either erupts onto the solid surface of the Earth or, if it does not erupt, cools to form intrusive igneous bodies. By crystallization of mafic magma, partial melting of solid rock and other processes, crust has been generated indirectly or directly from the mantle over geological time. Subject to geophysical constraints, all our real knowledge of the internal constitution of the Earth comes from the samples of the mantle represented by mafic magmatism (including blocks of solid mantle ripped off and carried up) and, indirectly, from the nature of the crust. For this reason, mafic magmatism throughout geological time, and especially in early Precambrian time (3.8–2 gigayears ago), forms the rock upon which our church of knowledge is built. That is the reason for such a book as this.

The notion of acid and basic in rocks and silicate melts has nothing to do with hydrogen ions, but stems instead from the old concept of silica (SiO_2) as silicic acid. Rocks poor in silica (and conversely, rich in magnesium-iron silicate) are, by default, basic. The trend away from this tradition among researchers in the United States, and towards the use of the acronyms felsic and mafic, evidently still awaits wide acceptance in Europe. An encouraging aspect of this entire collec-

tion of articles in *Early Precambrian Basic Magmatism* is the abundant demonstration that original rock chemistry can be recognized in the oldest rocks through a screen of variable (but quantifiable) alteration and metamorphic recrystallization. Students of modern volcanism who doubt this fact should take another look here.

The first part of the book deals with general aspects of early Precambrian magmatism, including that on the Moon. One wonders why the editors stopped there, when Mars and some meteorites could certainly qualify in this context. This part



Smoky atmosphere — after the initial horizontal eruption of Mount St Helens in 1980, a vertical explosion forced a column of hot gas, ash and rock 19 km into the atmosphere. *Our Mysterious Planet: Mysteries of the Natural World*, edited by Philip Whitfield, offers possible answers to questions concerning volcanoes, earthquakes, lightning, global climate, gravity, the demise of the dinosaurs and other of the Earth's enigmas. Published by Cassell, price is £18.95.

concludes with a chapter on mineralization associated with early Precambrian mafic magmatism, most notably the major (and classically strategic) deposits of nickel, chrome, gold and platinum group elements. The second part consists of regional syntheses, inevitably covering some of the same ground as in part I, but in some cases containing different points of view. There are strongly argued accounts of signally important occurrences in South Africa and Australia, and welcome chapters on little-known terranes in China and South America. Some may perceive too much detail on sites in the

United States and too little on those in Canada.

A main aim of studies such as those described here is the prospect of learning about the thermal history of the Earth, for the particular mineralogical and chemical characters of basic magmatism directly reflect the potential temperature of the upwelling source rocks. But there are serious impediments to making general statements about the bulk Earth through time, most of them reviewed carefully here by Bickle. There are unresolved paradoxes in the conflicting evidence from different parts of the crust. One such

paradox is ignored. At present, between 20 and 40 per cent (many favour the upper limit) of the heat coming from within the Earth is released by the freezing of the core. Most of the emphasis on the thermal history of the Earth dwells on the secure knowledge that internal heat production from radioactive decay has decreased by about three-fold since 3.5 gigayears ago. From this fact alone it can be concluded that the Earth now cools at a slower rate than in the past, and was hotter earlier. But core freezing could conceivably accelerate as radiogenic heat production decreases, in effect leading to a more-or-less constant release of heat over time. The record of the rocks evidently permits this model, but perhaps does not allow a clear choice among the alternatives. It is surprising not to find core heat brought centrally into the discussions in this volume.

What emerges, however, is a great wealth of information conveniently collected, but an equally great confusion about what it all means. The concluding remarks by the authors of most chapters serve as clear pointers to further study, but also, with distressing frequency, as statements of conflicting interpretations still unresolved even after decades of enlightened research in the field and laboratory. It is, perhaps, not surprising that this should be the case where the geological record is complex and incomplete but, in sum, it does suggest that the science is not yet up to the challenge of the rocks. Good news for the young. □

S. A. Morse is in the Department of Geology and Geography, University of Massachusetts, Amherst, Massachusetts 01003, USA.

Latitudinal distribution of solar-wind speed from magnetic observations of the Sun

Y.-M. Wang, N. R. Sheeley Jr & A. G. Nash

Code 4172, E O Hulburt Center for Space Research, Naval Research Laboratory, Washington DC 20375-5000, USA

Empirical studies suggest a close relationship between the solar-wind speed near the Earth and the magnetic structure of the solar corona. The correlation can be used to infer the latitudinal distribution of wind speed at different phases of the sunspot cycle, and to identify the sources of fast, high-latitude wind streams such as those that might be encountered by the Ulysses spacecraft on its journey toward the solar poles during 1992–1995.

PRESENT knowledge of the solar wind is largely based on spacecraft observations made near the ecliptic plane. The ecliptic wind shows large temporal and spatial variations, with bulk speeds ranging from 300–400 km s⁻¹ near the heliospheric current sheet to well over 700 km s⁻¹ in some high-speed streams¹. Long-term changes associated with the 11-yr sunspot cycle are also observed, with high-speed wind much more prevalent during the declining phase of the cycle than in the years near or preceding sunspot maximum². During the 1973–1974 Skylab mission, the occurrence of high-speed streams near the Earth and of the associated geomagnetic activity were found to accompany the passage of large, near-equatorial coronal holes across the solar disk³. Such studies led to the identification of coronal holes (regions of very low plasma density and open magnetic field) as sources of fast solar wind⁴.

Near sunspot minimum, the Sun's polar regions are dominated by large coronal holes which typically extend down to a latitude of 60° in each hemisphere. The oppositely directed magnetic field lines from the two polar holes meet at the heliospheric current sheet, which is relatively flat and lies close to the ecliptic plane at this time. The solar wind-speed profile has a minimum at the current sheet, but rises steeply on each side of it with an average latitudinal gradient of the order of 10 km s⁻¹ deg⁻¹ (refs 5, 6). With increased sunspot activity, the current sheet becomes warped and tilted with respect to the ecliptic plane; however, except near sunspot maximum, when the polar holes disappear, the general tendency for the wind speeds to increase with angular distance from the current sheet is found to persist^{5–8}. Such analyses, as well as recent Pioneer 11 observations^{9,10} at a heliographic latitude of ~16°, have suggested that wind speeds of 600 km s⁻¹ or more may occur persistently above the polar holes, and that high wind speeds may generally be far more prevalent at high latitudes than near the ecliptic.

The three-dimensional structure of the solar wind can be measured indirectly by the interplanetary scintillation (IPS) technique, which exploits the scattering of radio waves from distant sources by density inhomogeneities in the wind^{11–15}. Such observations have confirmed the presence of a high-speed polar wind, the boundaries of which approach the ecliptic near sunspot minimum but recede sharply toward sunspot maximum. Although it provides valuable information on the average properties of the global wind, the IPS method is limited by the small number of radio sources suitable for tracking, which results in poor temporal and spatial resolution. Further smearing is

caused by the line-of-sight integration of the signal; IPS measurements thus tend to underestimate the actual wind speeds and their latitudinal gradients^{13–15}.

Here we use a new procedure to deduce the structure and evolution of the out-of-ecliptic solar wind. The method relies on an empirical relation between daily wind speed near the Earth and the magnetic field in the solar corona, as determined by a current-free extrapolation of the observed photospheric field. The tendency for fast wind to be associated with slowly diverging coronal field was earlier noted by Levine *et al.*¹⁶ in a study based on a small number of coronal holes observed by Skylab. The hypothesis is supported by a recent analysis¹⁷ of observations of the daily wind speed and solar magnetic field extending over the past 22 years. Although derived from data on the ecliptic wind, the relationship should hold at any heliographic latitude or longitude, and we use the method here to infer the wind-speed evolution during 1976–1989 as a function of latitude.

The Ulysses mission¹⁸, scheduled for launch in October 1990, will obtain *in situ* measurements of the out-of-ecliptic solar wind after its Jupiter flyby in February 1992. Using solar magnetic data from the corresponding phase of the preceding sunspot cycle, we show that the spacecraft is likely to encounter very fast wind streams at mid-heliographic latitudes. The polar wind is predicted to attain its peak speeds of ~700 km s⁻¹ during 1991–1993, but to have slowed to ≤650 km s⁻¹ by the time Ulysses reaches its highest latitudes of ±70°–80° in 1994 and 1995.

The flux-tube expansion factor

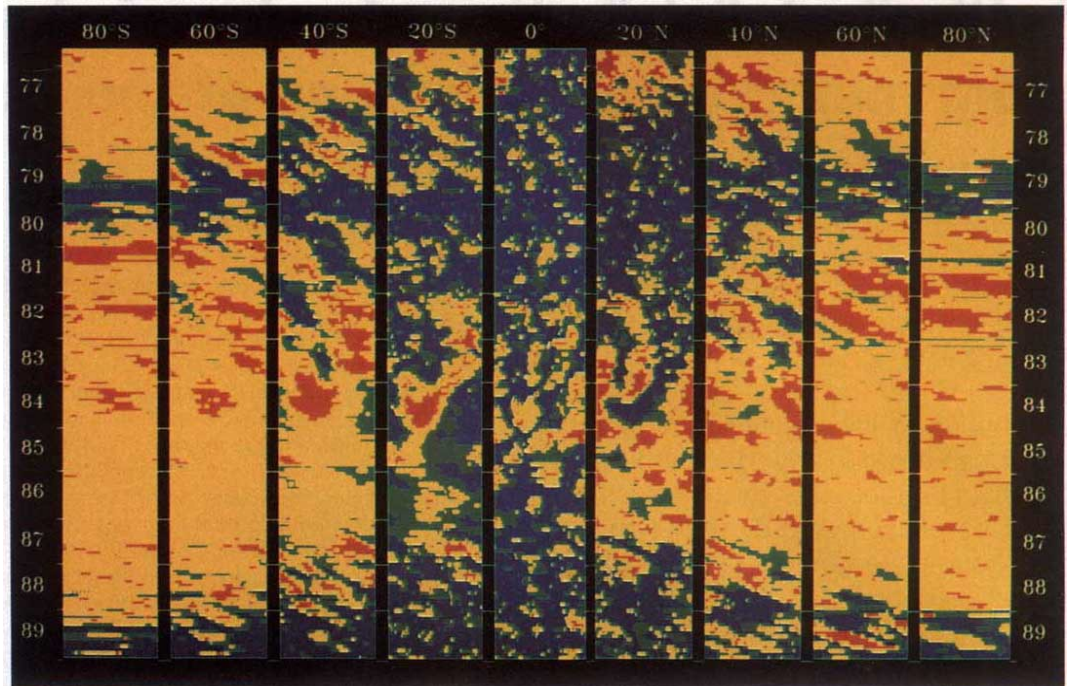
We infer the magnetic-field configuration of the solar corona using the potential-field source-surface model^{19,20}. We define spherical coordinates (r, θ, ϕ), where r is radial distance from the Sun's centre, θ is colatitude measured with respect to the solar rotation axis, and ϕ is Carrington longitude (increasing in the direction of the Sun's rotation). Electrical currents are neglected between the photosphere $r = R_\odot$ and a 'source surface' $r = R_s$, where the effect of the solar-wind flow is simulated by requiring the non-radial components of the magnetic field $B(r, \theta, \phi)$ to vanish. The further constraint that $B_r(R_\odot, \theta, \phi)$ match the observed photospheric flux distribution then allows the coefficients of the multipole expansion for B to be determined. By construction, field lines that cross the source surface are 'open', and their photospheric footpoint areas represent 'coronal holes'. The source-surface radius R_s is set to 2.5 $R_\odot$, a value consistent with the sizes of polar holes near sunspot minimum²¹ and with the polarity structure of the interplanetary magnetic field²².

For the photospheric magnetic observations, we use monthly synoptic maps provided by the Wilcox Solar Observatory (WSO), which have dimensions of 72 pixels in longitude by 30 pixels in sine latitude. The use of higher-resolution data from the Mount Wilson Observatory was found to yield qualitatively similar, but somewhat noisier, wind-speed patterns.

For any angular position (θ_s, ϕ_s) on the source surface, the flux-tube expansion factor f_s is defined as the ratio

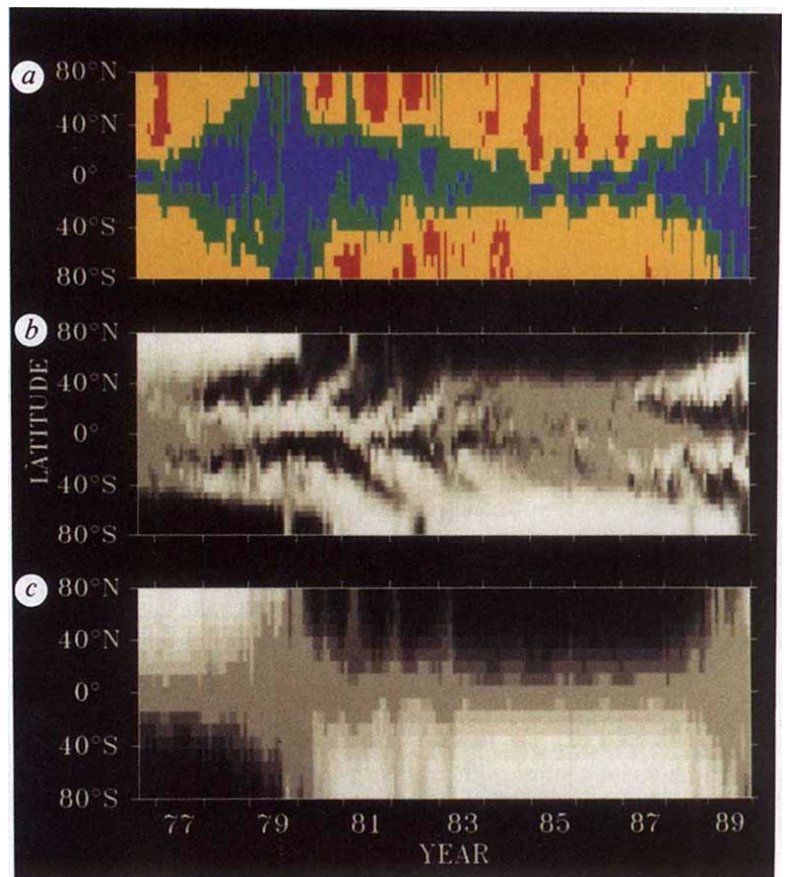
$$f_s(\theta_s, \phi_s) = \frac{R_\odot^2 B_r(R_\odot, \theta_\odot, \phi_\odot)}{R_s^2 B_r(R_s, \theta_s, \phi_s)} \quad (1)$$

FIG. 1 Evolution of solar-wind speed during 1976–1989 (Carrington rotations 1645–1823), for heliographic latitudes between 80°S and 80°N . In these plots, each row of coloured pixels represents a Carrington rotation, and Carrington longitude runs from right to left. (The format of each panel is similar to that of a Bartels plot, except that the 27-d Bartels rotations have been replaced by ~ 27.3 -d Carrington rotations.) Wind speeds are coded as follows: red ($v_w > 650 \text{ km s}^{-1}$); yellow ($550 \text{ km s}^{-1} < v_w \leq 650 \text{ km s}^{-1}$); green ($450 \text{ km s}^{-1} < v_w \leq 550 \text{ km s}^{-1}$); blue ($v_w \leq 450 \text{ km s}^{-1}$). A current-free approximation was used to extrapolate the observed WSO photospheric field to a source surface at $2.5 R_\odot$, where the distribution of flux-tube expansion factors was calculated and converted into wind speeds using the empirical scheme of ref. 17. In correcting the magnetic measurements for line-of-sight projection, we assume that the large-scale photospheric field is predominantly radial³⁵. The streaky appear-



ance of the wind-speed patterns in the highest-latitude panels reflects the fact that the polar fields are poorly observed at latitudes above $\sim 70^\circ$; the smaller-scale structures should therefore be regarded as noise.

FIG. 2 *a*, Longitudinally averaged wind speed $\bar{v}_w$ as a function of latitude and time. The averaging procedure is described in ref. 17. Colour coding: red ($\bar{v}_w > 630 \text{ km s}^{-1}$); yellow ($550 \text{ km s}^{-1} < \bar{v}_w \leq 630 \text{ km s}^{-1}$); green ($450 \text{ km s}^{-1} < \bar{v}_w \leq 550 \text{ km s}^{-1}$); blue ($\bar{v}_w \leq 450 \text{ km s}^{-1}$). *b*, Longitudinally averaged, radial photospheric field B_{ph} as a function of latitude and time. Grey-scale calibration: white ($B_{ph} > 5.5 \text{ G}$); neutral grey ($|B_{ph}| < 0.3 \text{ G}$); black ($B_{ph} < -5.5 \text{ G}$). *c*, Longitudinally averaged source-surface field B_{ss} at $2.5 R_\odot$ as a function of latitude and time. Grey-scale calibration: white ($B_{ss} > 0.3 \text{ G}$); neutral grey ($|B_{ss}| < 0.04 \text{ G}$); black ($B_{ss} < -0.3 \text{ G}$). Wind speeds and field strengths were determined using the WSO photospheric maps for Carrington rotations 1645–1823. The magnetograph measurements were multiplied by 1.8 to correct for the saturation of the Fe I 5,250 Å line profile²³. The annual modulation seen in the wind-speed structures during 1984–1987 is an artefact related to the Sun's 7.25° axial tilt, the residual effects of which are present in the low-resolution WSO data.



Here $(\theta_{\odot}, \phi_{\odot})$ are the footpoint coordinates of the field line that traverses (θ_s, ϕ_s) . The parameter f_s represents the factor by which an infinitesimal flux tube expands in solid angle between the photosphere and the source surface. In an earlier analysis¹⁷ of ecliptic wind measurements and magnetograph data during 1967–1988, we found an inverse correlation between the daily wind speed near the Earth, v_w , and the value of f_s at the corresponding sub-Earth location on the source surface (see also ref. 16). Ranges of v_w and f_s were related according to the following empirical scheme: $v_w > 650 \text{ km s}^{-1}$ ($f_s < 3.5$); $650 \text{ km s}^{-1} \geq v_w > 550 \text{ km s}^{-1}$ ($3.5 \leq f_s < 9$); $550 \text{ km s}^{-1} \geq v_w > 450 \text{ km s}^{-1}$ ($9 \leq f_s < 18$); $v_w \leq 450 \text{ km s}^{-1}$ ($f_s \geq 18$). Here we assume that the same correspondences apply to wind streams and flux tubes directed out of the ecliptic plane. We also suppose that the solar wind propagates radially outward from the source surface, so that the distribution of expansion factors at $2.5 R_{\odot}$ maps into that of wind speeds at 1 AU (the Earth's radial distance) without any latitudinal shift.

Latitudinal distribution during 1976–1989

Applying the above procedure to the WSO photospheric-field measurements, we have deduced the angular distribution of wind speeds at 1 AU during Carrington rotations 1645–1823, which span the interval between August 1976 and December 1989. Figure 1 displays the evolution of the wind speeds at a sequence of heliographic latitudes ranging from -80° to $+80^{\circ}$ in 20° intervals. In these plots, each row of pixels represents a single Carrington rotation (average duration 27.3 d), with Carrington longitude running from right to left. The longitudinal positions assigned to the wind speeds correspond to the time when the flow crossed the source surface (thus no correction was made for solar rotation during transit to 1 AU).

Figure 1 shows a number of striking properties of the global solar wind. (1) The wind tends to be much slower near the equator than at high latitudes. (2) The wind-speed structure is far more varied at low and middle latitudes, where a wide range of wind speeds often occurs during a given Carrington rotation, than near the poles (60° and above), where the wind speeds tend to be uniformly high. (3) At all latitudes, there is an interval around the 1979 and 1989 sunspot maxima when the wind speeds are reduced; this lull is briefest and most sharply defined at high latitudes. (4) At the equator, episodes of high-speed wind occur mainly during the declining phase of the cycle (1982–1985); at higher latitudes, the increasingly uniform distribution of high-speed wind spreads past sunspot minimum (1976 and 1986) toward the maximum phase of the next cycle. (5) Near the poles, the fastest wind streams (with $v_w > 650 \text{ km s}^{-1}$) occur just after sunspot maximum (1980–1982), not around sunspot minimum. (6) Through much of the rising and declining phases of the cycle, the fastest streams tend to occur at mid-latitudes, not near the poles.

Figure 2a shows the behaviour of the longitudinally averaged wind speeds as a function of latitude and time. For comparison, we have also plotted the longitudinally averaged photospheric and source-surface fields in latitude–time format (Fig. 2b, c). As shown by Fig. 2a, a band of low to moderately low ($\leq 550 \text{ km s}^{-1}$) average wind speed persists around the equator throughout the sunspot cycle. The band widens after sunspot minimum, spreading all the way to the poles by sunspot maximum; it then contracts, reaching its narrowest latitudinal extent of 10° – 20° on each side of the heliographic equator around sunspot minimum. Conversely, the regions of high-speed ($> 550 \text{ km s}^{-1}$) wind have their greatest latitudinal extent near sunspot minimum, but recede rapidly poleward as activity increases and vanish around sunspot maximum; they then abruptly re-form and gradually spread towards the equator during the declining phase of the cycle. Patches of very fast wind appear near the poles just after sunspot maximum, but their ‘centres of mass’ subsequently migrate to lower latitudes.

The latitude–time plot of the photospheric field (Fig. 2b)

shows the polar fields reversing around 1980 and 1989, and magnetic flux surging poleward from the sunspot latitudes during 1980–1983 and 1989. The series of alternating-polarity surges in each hemisphere gives rise to large fluctuations in the polar field just after sunspot maximum. The vanishing of the high-speed regions in Fig. 2a precedes the dissolution of the old-cycle polar fields by about a year, whereas the sudden reappearance of these regions and the generation of very fast polar wind streams coincide with the arrival of the surges at the poles.

The evolution of the polar fields is also reflected in the behaviour of the source-surface field, the axisymmetric component of which, B_{ss} , weakens and undergoes polarity reversal in each hemisphere near sunspot maximum (Fig. 2c). Comparison of Fig. 2a, c shows that the slow wind is centred about the latitudes where B_{ss} is weakest; this ‘neutral belt’ of the axisymmetric field lies close to the equator except near sunspot maximum. On the other hand, very fast wind sometimes extends into the regions of relatively weak B_{ss} at low latitudes.

Sources of fast and slow wind

In our empirical model, the entire solar wind (as well as the interplanetary field) originates from open magnetic regions on the Sun which we identify with coronal holes; the expansion rate of the coronal field lines determines whether the wind is fast or slow. In the years around sunspot minimum, the polar holes provide most of the Sun's open flux^{23,24}. The interiors of these large holes are characterized by relatively low flux-tube expansion factors; typically $f_s \approx 4$ – 5 near the polar axis, corresponding to asymptotic wind speeds of 600 – 650 km s^{-1} . The value of f_s rises steeply near the edge of each hole, however, where the field lines map into the low-latitude region around the source-surface neutral line; this ‘boundary layer’ behaviour gives rise to the narrow band of slow wind around the equator near sunspot minimum. With the onset of new-cycle sunspot activity and the emergence of strong bipolar magnetic regions at mid-latitudes, the broad unipolar areas centred around the poles and their embedded holes recede. The shrinking of the polar holes is accompanied by an increase in the flux-tube expansion rates, because the open field lines now fan out from a smaller area on the photosphere. Thus the band of low average wind speed spreads to higher latitudes, reaching the poles at sunspot maximum, when the polar holes have disappeared. At this time, the solar wind emanates from a collection of small, often elongated holes dispersed over a wide range of latitudes. The vanishing of the polar holes precedes the actual reversal of the polar fields, which occurs ~ 1 – 2 yr after sunspot maximum.

The evolution of the polar fields during the sunspot cycle may be understood in terms of the magnetic properties of active regions and the transport of magnetic flux across latitudes by supergranular convective motions and a meridional surface flow²⁵. Bipolar magnetic regions normally erupt with their dipole moments directed opposite to that of the old-cycle polar field. The lower-latitude, leading-polarity flux, which is of opposite sign in each hemisphere, is preferentially annihilated by supergranular ‘turbulent diffusion’ across the equator. This leaves a surplus of trailing-polarity flux in each hemisphere, which is carried poleward by meridional flow and reverses the polar field. The size of the polar holes near sunspot minimum depends on the degree of concentration of the polar fields, which in turn is determined by a balance between poleward flow and equatorward diffusion^{21,26}.

The fastest wind streams seen in Fig. 1 generally do not originate at the centres of the polar holes, where the field attains a local maximum, but around minimum-field regions separating neighbouring coronal holes of the same polarity²⁷. Unusually low values of f_s are associated with such configurations because the field lines that fan out from the two holes meet along their ‘symmetry plane’ and are deflected into a more radial direction; as a result, the facing sides of the holes become sources of very fast wind.

In practice, slowly diverging flux tubes are often rooted in the weak-field 'gaps' between the flux concentrations in the sunspot belts and the polar caps. As active regions decay, coronal holes form in the spreading remnants and interact with (and sometimes become attached to) the like-polarity polar holes. Thus, during the rising phase of the sunspot cycle, very fast wind originates from the poleward sides of the sheared holes associated with mid-latitude active regions. Such wind streams are prominent around $\pm 40^\circ$ during 1977–1979 and 1987–1988 (Fig. 1). The slanted orientation of the patterns reflects the Sun's differential rotation: the mid-latitude holes rotate relatively slowly and thus drift backward in the 27.3-d Carrington frame. During the declining phase of the cycle, very fast wind emanates from the rigidly rotating extensions that link the polar holes with decaying low-latitude active regions. These streams may cover a wide range of latitudes and often extend into the ecliptic.

A different type of minimum-field configuration is created at the Sun's poles just after sunspot maximum, when the old-cycle polar fields have been neutralized. At this time, waves of unipolar flux from the sunspot belts arrive at the poles (see Fig. 2b) and begin to establish the new-cycle polar fields. The surges, which are apparently triggered by an acceleration of the meridional flow rate²⁵, carry east-west elongated holes. As the holes converge onto the weak-field regions around the poles, their field lines 'collide' and are deflected outward along the polar axis. The resulting flux-tube expansion rates are lower than those occurring above the poles later in the cycle, when the polar fields have been fully established. Thus the polar wind is faster just after sunspot maximum than around sunspot minimum.

A 'preview' of the Ulysses mission

The Ulysses spacecraft will perform *in situ* observations of the solar wind at heliographic latitudes of up to $\pm 80^\circ$ during the declining years of the current sunspot cycle¹⁸. After leaving the ecliptic near Jupiter in February 1992, it will attain its southern-

most latitudes (at a heliocentric distance of ~ 2 AU) during 1994, cross back through the ecliptic plane, and ascend to its northernmost latitudes in 1995.

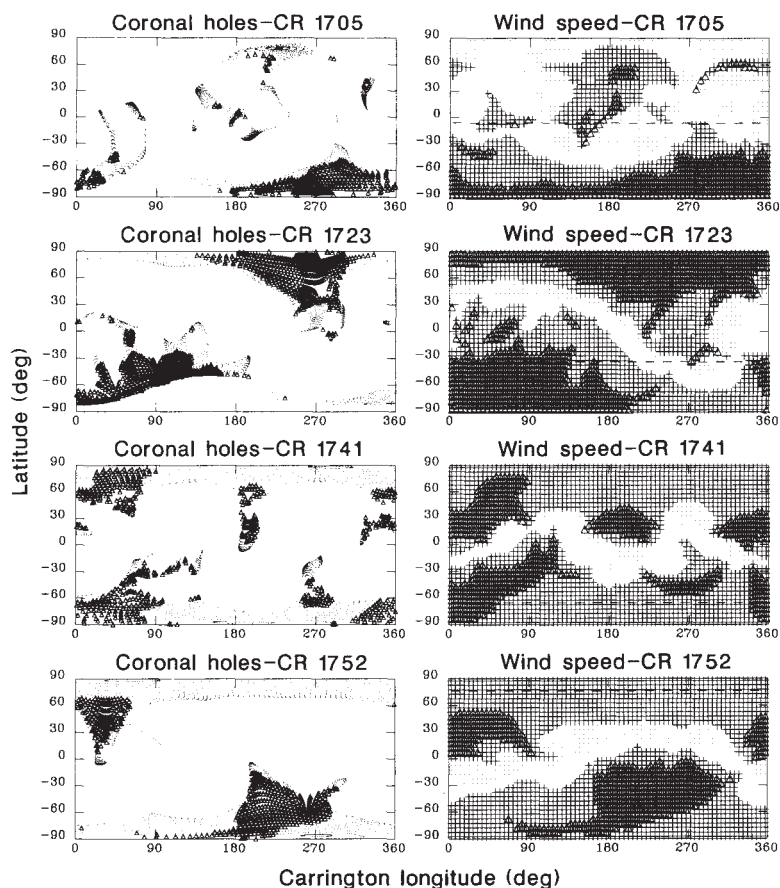
Below we describe the wind-speed structures that a similar mission might have recorded 11 years earlier, during the declining phase of sunspot cycle 21. Although the general trends are apparent from Fig. 1, we examine in detail a representative sample of Carrington rotations (1705, 1723, 1741 and 1752) from the period 1981–1984. Using the corresponding WSO synoptic maps, we have determined the configuration of coronal holes as well as the angular distribution of wind speeds during these rotations (Fig. 3). We consider each Carrington rotation in turn.

CR 1705 (February 1981). A giant surge of positive-polarity flux carrying a large east-west hole is now reaching the Sun's south pole. The field lines from the poleward side of the hole converge above the weak, closed-field area at the pole, producing a very fast polar 'jet'. Narrow trans-equatorial holes generate small patches of high-speed wind near the ecliptic; again, the fastest streams originate around minimum-field regions within these fragmentary holes.

CR 1723 (June 1982). The north pole is now covered by a negative-polarity (new cycle) hole. Very fast wind is centred around the 'gaps' within the lobe of this asymmetric polar hole. The arrival of a negative-polarity surge at the south pole has set back the development of the south polar hole. But a counter-surge carrying a large positive-polarity hole, with fast wind issuing from its poleward side, is about to re-establish the new-cycle polar field.

CR 1741 (October 1983). Both polar holes are now fully established, and extend to a latitude of $\pm 60^\circ$. In addition, a number of lower-latitude holes have formed in the remnants of active regions. As the active-region flux spreads poleward, the embedded holes gradually become attached to the like-polarity polar holes. The resulting extensions rotate quasi-rigidly at the rate of the remnant low-latitude flux²⁸. The fastest streams

FIG. 3 Distribution of coronal holes at the photosphere and wind speeds at 1 AU during Carrington rotations 1705, 1723, 1741 and 1752 (declining phase of sunspot cycle 21). Open field regions ('coronal holes'), indicated by the dotted areas in the left-hand panels, were determined by tracing field lines from the source surface at $2.5 R_\odot$ down to the photosphere. Triangles mark the sources of very fast wind ($v_w > 650 \text{ km s}^{-1}$). The wind speeds, inferred from the flux-tube expansion factors f_s and shown in the right-hand panels, are coded as follows: Δ , $v_w > 650 \text{ km s}^{-1}$; +, $550 \text{ km s}^{-1} < v_w \leq 650 \text{ km s}^{-1}$. To indicate the location of the magnetic neutral line, dots are plotted wherever $v_w \leq 550 \text{ km s}^{-1}$ and the source-surface field has negative polarity. The longitudes assigned to the wind speeds refer to the time when the flow crossed the source surface, and have not been shifted to account for the solar rotation during transit to 1 AU. The dashed line in each wind-speed map shows the expected latitudinal location of the Ulysses spacecraft (at 1–5 AU) 11 years after the starting date of the Carrington rotation.



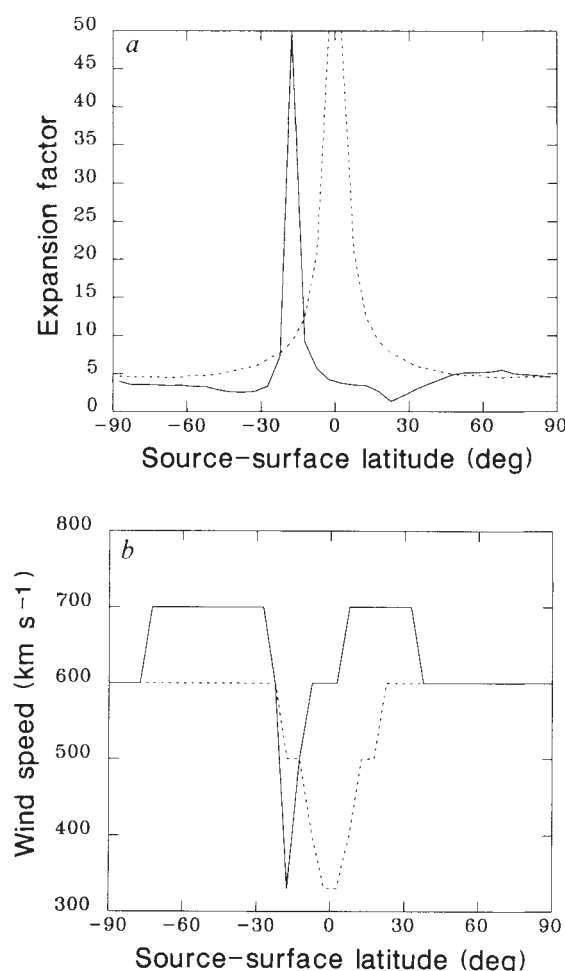


FIG. 4 *a*, Solid line shows the flux-tube expansion factor f_s near longitude 340° during CR 1741 as a function of source-surface latitude. Also plotted (dotted line) is the latitudinal variation of f_s corresponding to a photospheric field distribution of the form $\pm \cos^8 \theta$, for which the polar holes are axisymmetric and extend to latitude 60°. For the same two cases, *b* shows the corresponding wind speeds as a function of latitude, as inferred using Table 2 of ref. 17.

present at this time are centred around the weak-field areas between the low-latitude and polar holes.

CR 1752 (August 1984). With sunspot activity continuing to decline, each polar hole now shows a single large lobe, and the source-surface field has a corresponding tilted-dipole structure. But, the fastest wind streams, which are associated with the minimum-field regions within the lobes, are not centred around the dipole axis but are displaced to lower latitudes.

Latitudinal gradients of wind speed

The wind-speed maps of Fig. 3 suggest the presence of large gradients near the boundaries of high-speed streams, particularly in the vicinity of the magnetic neutral line. As an example taken from CR 1741, we have plotted the latitudinal variation of f_s and the associated wind speeds near Carrington longitude 340° (Fig. 4). Over a latitudinal distance of only 5° on each side of the neutral line (located near latitude 17.5° S), the flux-tube expansion factor decreases from infinity to ≤ 9 ; the corresponding wind-speed gradients are of the order of $50 \text{ km s}^{-1} \text{ deg}^{-1}$. Local latitudinal gradients of this magnitude have been inferred from multi-spacecraft observations near the ecliptic during both the rising²⁹ and declining³⁰ phases of sunspot cycle 20. In March 1975, a gradient of at least $30 \text{ km s}^{-1} \text{ deg}^{-1}$ was found at the low-latitude edge of a high-speed stream from a polar-hole extension³⁰.

Figure 4 also shows the latitudinal variation of f_s and inferred wind speed for a photospheric flux distribution of the form $\pm \cos^8 \theta$, which approximates the profile of the longitudinally averaged field near sunspot minimum^{23–26}. In this case the polar holes are axisymmetric and extend down to latitude $\pm 60^\circ$, and the neutral line of the source-surface field coincides with the heliographic equator. The value of f_s now decreases from infinity at the equator to ~ 9 at latitude $\pm 17.5^\circ$, corresponding to an increase of the wind speed to $\sim 550 \text{ km s}^{-1}$ within 17.5° of the neutral line. Spacecraft measurements show average wind-speed gradients of similar magnitude ($\sim 10 \text{ km s}^{-1} \text{ deg}^{-1}$) on each side of the heliospheric current sheet around sunspot minimum^{5,6,9,10}.

The model thus accounts for the observation of both extremely steep local gradients and moderately steep average gradients in the low-latitude wind, with the former occurring at the boundaries of very fast streams from the polar-hole extensions, and the latter associated with the boundaries of the moderately fast, persistent wind from the polar holes themselves.

Concluding remarks

This study suggests that the solar wind outside the ecliptic plane has a variable and non-uniform structure, particularly at latitudes below $\sim 40^\circ$. Very fast streams (with speeds of the order of 700 km s^{-1}) are generated near the poles just after sunspot maximum, but their centroidal location migrates to lower latitudes during the declining phase of the sunspot cycle. Thus as Ulysses journeys away from the ecliptic during 1992–1994, it will tend to encounter even faster wind at mid-latitudes than at its destination above the Sun's poles.

It has frequently been suggested that the solar wind is organised in a simple way around the heliospheric current sheet, with the wind speed increasing progressively with angular distance from this 'heliomagnetic equator'^{7,8}. According to this view, high-speed streams are observed at Earth when the heliomagnetic coordinate system becomes significantly tilted with respect to the ecliptic. We have found, however, that the wind speeds generated by the interaction between an active-region remnant and the adjacent like-polarity polar hole exceed those occurring near the axis of an unperturbed polar hole. Thus the recurrent high-speed streams observed late in the sunspot cycle do not result simply from the 'tipping' of the polar-hole axis, but are intrinsic to the weak-field extensions; the fastest wind is in fact centred around the warps in the heliospheric current sheet.

The physical basis for the relationship between solar-wind speed and coronal flux-tube expansion remains to be understood, and its implications for the mechanism of solar-wind acceleration need to be explored. Simple models in which thermal or Alfvén-wave pressure gradients provide the acceleration yield only a weak dependence of the asymptotic wind speeds on the expansion factor³¹. In such calculations, the magnetic-field structure as well as the coronal base temperature and density are arbitrarily prescribed. An understanding of the part played by the flux-tube expansion factor may require a magnetohydrodynamic treatment in which the flow is allowed to react back on the magnetic field³². It has also been suggested that mechanical energy may be preferentially dissipated along rapidly expanding flux tubes, before it can be deposited beyond the sonic point where it drives the wind most efficiently^{33,34}. It remains to be seen whether the development of realistic magnetohydrodynamic models, perhaps coupled with an improved description of wave propagation in the corona, will lead to a better understanding of solar-wind acceleration. □

Received 23 May; accepted 7 September 1990.

1. Feldman, W. C., Asbridge, J. R., Bame, S. J. & Gosling, J. T. *J. geophys. Res.* **83**, 2177–2189 (1978).
2. Sheeley, N. R. Jr, Asbridge, J. R., Bame, S. J. & Harvey, J. W. *Solar Phys.* **52**, 485–495 (1977).
3. Nolte, J. T. et al. *Solar Phys.* **46**, 303–322 (1976).
4. Zirker, J. B. *Coronal Holes and High Speed Wind Streams* (Colorado Associated University Press, Boulder, 1977).

5. Newkirk, G. Jr & Fisk, L. A. *J. geophys. Res.* **90**, 3391-3414 (1985).
6. Bruno, R., Villante, U., Bavassano, B., Schwenn, R. & Mariani, F. *Solar Phys.* **104**, 431-445 (1986).
7. Hakamada, K. & Akasofu, S.-I. *J. geophys. Res.* **86**, 1290-1298 (1981).
8. Zhao, X.-P. & Hundhausen, A. J. *J. geophys. Res.* **86**, 5423-5430 (1981).
9. Gazis, P. R., Mihalov, J. D., Barnes, A., Lazarus, A. J. & Smith, E. J. *Geophys. Res. Lett.* **16**, 223-226 (1989).
10. Mihalov, J. D., Barnes, A., Hundhausen, A. J. & Smith, E. J. *J. geophys. Res.* **95**, 8231-8242 (1990).
11. Dennison, P. A. & Hewish, A. *Nature* **213**, 343-346 (1967).
12. Coles, W. A. *et al. Nature* **286**, 239-241 (1980).
13. Kojima, M. & Kakinuma, T. *J. geophys. Res.* **92**, 7269-7279 (1987).
14. Watanabe, T. *Adv. Space Res.* **9**, 99-110 (1989).
15. Coles, W. A., Harmon, J. K., Lazarus, A. J. & Sullivan, J. D. *J. geophys. Res.* **83**, 3337-3341 (1978).
16. Levine, R. H., Altschuler, M. D. & Harvey, J. W. *J. geophys. Res.* **82**, 1061-1065 (1977).
17. Wang, Y.-M. & Sheeley, N. R. Jr *Astrophys. J.* **355**, 726-732 (1990).
18. Wenzel, K.-P., Marsden, R. G., Page, D. E. & Smith, E. J. *Adv. Space Res.* **9**, 25-29 (1989).
19. Schatten, K. H., Wilcox, J. M. & Ness, N. F. *Solar Phys.* **6**, 442-455 (1969).
20. Altschuler, M. D. & Newkirk, G. *Solar Phys.* **9**, 131-149 (1969).
21. Sheeley, N. R. Jr, Wang, Y.-M. & Harvey, J. W. *Solar Phys.* **119**, 323-340 (1989).
22. Hoeksema, J. T. thesis, Stanford Univ. (1984).
23. Svalgaard, L., Duvall, T. L. Jr & Scherrer, P. H. *Solar Phys.* **58**, 225-240 (1978).
24. Wang, Y.-M. & Sheeley, N. R. Jr *J. geophys. Res.* **93**, 11227-11236 (1988).
25. Wang, Y.-M., Nash, A. G. & Sheeley, N. R. Jr *Astrophys. J.* **347**, 529-539 (1989).
26. Sheeley, N. R. Jr, Wang, Y.-M. & DeVore, C. R. *Solar Phys.* **124**, 1-13 (1989).
27. Sheeley, N. R. Jr & Wang, Y.-M. *Solar Phys.* (in the press).
28. Nash, A. G., Sheeley, N. R. Jr & Wang, Y.-M. *Solar Phys.* **117**, 359-389 (1988).
29. Rhodes, E. J. Jr & Smith, E. J. *J. geophys. Res.* **86**, 8877-8892 (1981).
30. Schwenn, R. *et al. J. geophys. Res.* **83**, 1011-1017 (1978).
31. Leer, E., Holzer, T. E. & Flå, T. *Space Sci. Rev.* **33**, 161-200 (1982).
32. Fichtner, H. & Fahr, H. J. *Planet. Space Sci.* **37**, 987-999 (1989).
33. Rosner, R. & Vaiana, G. S. *Astrophys. J.* **216**, 141-157 (1977).
34. Withbroe, G. L. *Astrophys. J.* **325**, 442-467 (1988).
35. Howard, R. & LaBonte, B. J. *Solar Phys.* **74**, 131-145 (1981).

ACKNOWLEDGEMENTS. We thank J. T. Hoeksema and P. H. Scherrer for supplying the WSO magnetograph measurements, and E. J. Smith for providing information about the orbital trajectory of Ulysses. This work was supported by NASA and by the Office of Naval Research.

Identification of a receptor for protein import into mitochondria

Debkumar Pain, Hiroshi Murakami & Günter Blobel

Laboratory of Cell Biology, Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10021, USA

Anti-idiotypic antibodies, prepared using a chemically synthesized signal peptide of a mitochondrial precursor protein, recognized a mitochondrial integral membrane protein (p32). Fab fragments derived from both anti-idiotypic antibodies and monospecific antibodies against purified p32 inhibited protein import into mitochondria. Moreover, anti-p32 antibodies specifically immunoprecipitated a precursor-p32 complex after detergent solubilization of mitochondria. Immunoelectron microscopy and subfractionation of mitochondria indicate that p32 is located in contact sites between the outer and inner mitochondrial membranes.

MOST mitochondrial proteins are synthesized as large precursors¹ containing an N-terminal signal sequence^{2,3} that targets these proteins for import into mitochondria. Protein import occurs at contact sites between the inner and outer mitochondrial membranes⁴, and is dependent on protein components of both the cytoplasm (see ref. 5) and the mitochondrial membranes⁶⁻¹⁰. Here we report the results of an anti-idiotypic antibody approach to identify mitochondrial protein component(s) that interact with the signal sequence.

Anti-idiotypic antibodies inhibit import

The strategy for the identification of protein(s) that interact with the mitochondrial signal sequence was similar to that previously used to identify a chloroplast import receptor¹¹. The signal peptide¹² of the precursor of subunit IV of cytochrome *c* oxidase (pCoxIV) was chemically synthesized, and antibodies against it were produced in rabbits. These antibodies immunoprecipitated *in vitro* synthesized pCoxIV-DHFR, a chimaeric protein consisting of the N-terminal 22 residues of the signal sequence of pCoxIV linked to mouse dihydrofolate reductase (DHFR)¹³ (data not shown). A second set of rabbit antibodies was then produced against an IgG fraction of the anti-signal sequence antiserum to obtain anti-idiotypic antibodies (anti-IT) which, as expected, did not immunoprecipitate pCoxIV-DHFR (data not shown). Among these anti-IT antibodies are likely to be

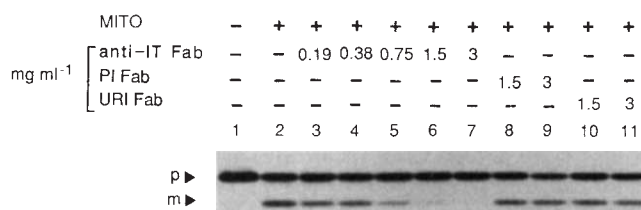


FIG. 1 Effect of monovalent Fab fragments of various IgGs on *in vitro* import of precursor into yeast mitochondria. A post-ribosomal supernatant of a reticulocyte lysate translation mixture containing newly-synthesized precursor (pCoxIV-DHFR) was incubated in a total volume of 0.15 ml for 6 min at 30 °C with isolated yeast mitochondria (50 µg protein) preincubated in the absence (lane 2) or the presence of various amounts (28.5-450 µg) of Fab fragments (lanes 3-11). Anti-IT, anti-idiotypic; PI, preimmune; URI, unrelated immune. Numbers indicate final concentrations of Fab (mg ml⁻¹). Lane 1 shows 25% of post-ribosomal supernatant used in the import assay. p, Precursor (pCoxIV-DHFR); m, mature (DHFR) forms.

METHODS. The plasmid pDS5/2-1-CoxIV-DHFR¹³ codes for a hybrid protein in which most of the cleavable signal sequence of the precursor of yeast cytochrome *c* oxidase subunit IV (22 out of 25 amino acids) is attached to the N terminus of mouse DHFR, a cytosolic protein. The *Pst*I-*Hind*III fragment of the plasmid was subcloned into the multiple cloning site of pGEM-3 vector (Promega) under the T7 promoter. The resulting plasmid pGT-CoxIV-DHFR was purified, linearized downstream from the gene with *Hind*III and transcribed using T7 RNA polymerase^{11,17}. The transcript was translated using a rabbit reticulocyte lysate cell-free translation system^{18,19}. Post-ribosomal supernatant containing the newly synthesized pCoxIV-DHFR was obtained by centrifuging the translation mixture for 22 min at 4 °C in a Beckman Airfuge at 140,000g. A 25 residue-long peptide representing the entire signal sequence¹² of cytochrome *c* oxidase subunit IV was chemically synthesized, and antibodies were produced in rabbits as previously described¹¹. Anti-idiotypic antiserum was generated in rabbits against an IgG fraction of the anti-signal sequence antiserum. IgG fractions from various decomplexed (56 °C for 20 min) rabbit sera (anti-IT, anti-idiotypic; PI, preimmune; URI, unrelated immune sera) were isolated and used to prepare monovalent Fab fragments as described¹¹. Mitochondria were isolated from *Saccharomyces cerevisiae* strain D273-10B (ATCC25657)⁵. Mitochondria (50 µg protein) were preincubated in the absence or presence of various Fab fragments at 4 °C for 30 min and the import reaction was carried out for 6 min at 30 °C essentially as described^{5,19}. Mitochondrial pellets, obtained after centrifugation at 4 °C in a microfuge (14,000g) for 10 min, were analysed by SDS-PAGE and the gels were fluorographed^{5,11,17}. Radioactivity in p and m forms in dried gels was measured with an AMBIS β scanner.

antibodies that mimic the signal peptide. These would be expected to bind to the signal sequence-binding domain of cytoplasmic signal recognition factor(s) or of membrane-integrated import receptor(s) and thereby inhibit import of precursors into mitochondria.

We prepared Fab fragments from anti-idiotypic IgG (anti-IT Fab) as well as from IgGs of preimmune serum (PI Fab) and an unrelated immune serum (URI Fab). These Fab fragments were then tested in an import reaction (Fig. 1), consisting of isolated yeast mitochondria supplemented with post-ribosomal supernatant prepared from a rabbit reticulocyte lysate cell-free system after translation of messenger RNA for pCoxIV-DHFR. Following incubation for import, mitochondria were sedimented and analysed by SDS-PAGE and fluorography. Some of the pCoxIV-DHFR polypeptides (p) that were present in the post-ribosomal supernatant (lane 1) were converted to mature DHFR (m) (Fig. 1, lane 2) as a result of import into yeast mitochondria and cleavage of the signal peptide by the matrix signal peptidase^{6,13}. When increasing concentrations of anti-IT Fab fragments were present in the import reaction, there was a concentration-dependent inhibition of import (see decrease in the number of mature polypeptides in lanes 3–7) with >95% inhibition of import at the highest concentrations of anti-IT Fab (lanes 6 and 7). In contrast, equivalent concentrations of preimmune Fab fragments (lanes 8 and 9) or unrelated immune Fab fragments (lanes 10 and 11) had only a slight effect. These data suggest that the anti-IT Fab fragments are indeed directed against the signal sequence-binding site of mitochondrial import receptor(s) and inhibit import by competing with the signal sequence for access to the signal sequence-binding site of the receptor(s).

FIG. 2 Identification and characterization of mitochondrial proteins reactive with anti-idiotypic antibodies. **a**, Proteins of either total mitochondria (lanes T) or of mitochondrial subfractions (lanes S and P) were separated by SDS-PAGE, transferred onto nitrocellulose and subsequently probed with anti-idiotypic antibodies (anti-IT) or preimmune IgG (PI) as indicated. Lane 1, M_r standards labelled with ¹⁴C. Lanes 2 and 3, pellet (P) and supernatant (S) fractions, respectively, of 100 μ g of mitochondria extracted with alkali (pH 11.5). Lanes 4 and 5, P and S fractions, respectively, of 100 μ g of mitochondria extracted with 0.5 M potassium acetate. Lanes 6 and 7, 100 μ g of unextracted mitochondria (T). **b**, Mitochondrial proteins (75 μ g) were separated by SDS-PAGE, transferred onto nitrocellulose and subsequently probed with anti-IT antibodies in the presence of either uncoupled BSA (lane 1) or BSA conjugated to the synthetic signal peptide of cytochrome *c* oxidase subunit IV (signal-BSA, 3–12 units ml^{-1} , lanes 2–4). **c**, Quantitation of the anti-IT antibody binding to p65/63 and p32 as determined by densitometric scanning of the immunoblot shown in **b**.

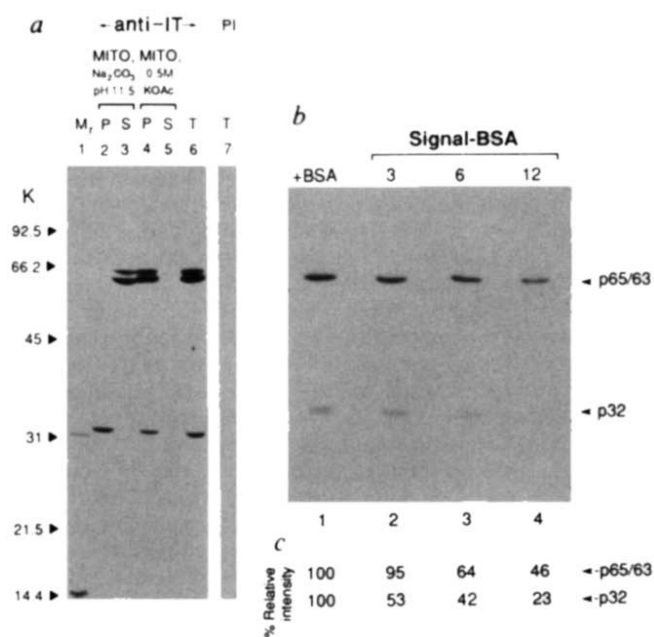
METHODS. **a**, Mitochondrial pellets (1 mg protein) were resuspended in 200 μ l of either 20 mM HEPES/KOH, pH 7.5, 0.5 M CH_3COOK or 0.1 M Na_2CO_3 , pH 11.5, both containing 5 mM EDTA, 1 mM phenylmethylsulphonyl fluoride, 50 units ml^{-1} trypsin and 5 μ g ml^{-1} each of antipain, chymostatin, leupeptin and pepstatin. The suspensions were vigorously mixed, and incubated on ice for 15 min, loaded onto a 0.8 ml cushion containing either 20 mM HEPES/KOH, pH 7.5, 0.25 M sucrose, 0.5 M CH_3COOK , 5 mM EDTA or 0.1 M Na_2CO_3 , pH 11.5, 0.25 M sucrose, 5 mM EDTA, and centrifuged at 356,000g for 20 min using the Beckman TLA-100.2 fixed angle rotor. Equivalent supernatant and pellet fractions were analysed by SDS-PAGE, transferred onto nitrocellulose, stained with amido black, destained, blocked with BSA and probed either with anti-IT IgG (10 mg ml^{-1}) or preimmune IgG (10 mg ml^{-1}) at a dilution of 1:500 by procedures described previously¹¹. **b**, The signal peptide (7.5 mg) was coupled to BSA (7.5 mg) using the water-soluble crosslinker 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide²⁰ and dialysed against appropriate buffer. BSA concentration in the dialysate (final volume 6.8 ml) was calculated to be 1.1 mg ml^{-1} . The resulting conjugate (signal-BSA) inhibited import of pCoxIV-DHFR into mitochondria in a dose-dependent manner (data not shown). One unit of the conjugate is defined as the amount necessary to block 50% import of pCoxIV-DHFR into mitochondria, using the assay conditions described in Fig. 1 methods, except that import was carried out for 15 min (instead of 6 min). One unit of signal-BSA conjugate contained 16.65 μ g BSA. Mitochondrial proteins (75 μ g per lane) were separated by SDS-PAGE on a 12% polyacrylamide gel, blotted onto nitrocellulose and blocked with 3% BSA in PBS as

As observed by others⁶, a large proportion of precursor molecules cosedimented with mitochondria (lanes 2–11), and it has been suggested⁶ that this cosedimentation is largely due to nonspecific and nonproductive binding. It seems that this binding is not inhibited by anti-IT Fab fragments (lanes 3–7) suggesting that precursors cosedimenting with mitochondria are unlikely to represent import receptor-bound translocation intermediates.

Anti-idiotypic antibody-reactive proteins

Isolated yeast mitochondria (Fig. 2a, lanes 6 and 7), as well as mitochondria extracted with 0.5 M potassium acetate (Fig. 2a, lanes 4 and 5) or Na_2CO_3 at pH 11.5 (Fig. 2a, lanes 2 and 3), were separated by SDS-PAGE, transferred to nitrocellulose and probed with the anti-IT antibodies (lanes 2–6) or the corresponding preimmune serum (lane 7). The anti-IT antibodies reacted with three proteins of estimated relative molecular mass 65,000, 63,000 (p65/63) and 32,000 (p32) (M_r 65K, 63K and 32K) when whole mitochondria were used (lane 6). None of the mitochondrial proteins reacted with the preimmune serum (lane 7). All three anti-IT antibody-reactive proteins remained associated with mitochondria after extraction with 0.5 M potassium acetate (lanes 4 and 5). However at pH 11.5, p65/63 were extracted, whereas p32 remained associated with the mitochondrial membrane (lanes 2 and 3). These data suggest that p65/63 are peripheral membrane proteins and that p32 is an integral membrane protein.

If the anti-IT antibodies mimic the signal sequence and if the immunoreactivity of p65/63 and p32 with these antibodies arises by virtue of a signal sequence-binding site, then reactivity with anti-IT antibodies should be inhibited by the signal peptide. To



described¹¹. The blots were then washed several times with buffer A (PBS containing 0.1% Triton X-100 and 0.2 mg ml^{-1} BSA) and immersed either in buffer A or buffer A containing 3–12 units ml^{-1} signal-BSA. The final BSA concentration in each case was adjusted to 0.2 mg ml^{-1} . Following incubation for 1 h at room temperature, anti-IT IgG (10 mg ml^{-1}) was added to a final dilution of 1:1,000 and incubated further for 3 h at room temperature. The blots were then washed first with buffer A and then with PBS containing 0.1% Triton X-100, 0.02% SDS and 0.1% BSA, and subsequently probed with ¹²⁵I-labelled Protein A as described¹¹. The p65 and p63 proteins were not completely resolved in the immunoblots in **b**, and therefore appeared as a single band. **c**, The relative intensity of p65/63 and p32 in each lane of **b** was determined by an ULTROSAN XL enhanced laser densitometer (Pharmacia LKB).

TABLE 1 Quantitation of gold particle distribution on mitochondrial surface

Experiments	Total no.	Distribution of gold particles				<i>P</i>
		CS	OM	CS/OM	CS + CS/OM	
1. anti-IT						
<i>O</i> (observed)	113	105	5	3	108	<0.001
<i>E</i> (expected)		(18)	(69.4)	(25.6)	(43.6)	
Distribution of gold particles (%)	100	93	4.4	2.6	95.6	
2. anti-p32						
<i>O</i>	132	117	8	7	124	<0.001
<i>E</i>		(24)	(83.3)	(24.7)	(48.7)	
Distribution of gold particles (%)	100	88.6	6.1	5.3	93.9	

Immunoelectron microscopy with anti-IT (experiment 1) and anti-p32 (experiment 2) is described in Fig. 6 methods. Morphometric analysis of bound gold particles on mitochondrial surface was essentially as described⁴. The regions of the mitochondrial outer membrane and the contact site were measured from electron micrographs at a linear magnification of 82,000. Regions (≤ 50 nm in length along the membrane) where the distance between centres of the inner and outer mitochondrial membranes was < 20 nm were considered as contact sites (CS). Regions where contacts between the outer and inner membranes extended > 50 nm were not considered to be contact sites as such. Instead, for statistical analysis these regions (CS/OM) were considered to represent a combination of both outer membrane (OM) and contact sites (CS). Gold particles at a distance of ≤ 30 nm from the outer membrane were considered membrane-associated and counted. In experiment 1, gold particles were counted on a total length of 62.7 μ m containing 10 μ m CS, 38.5 μ m OM and 14.2 μ m CS/OM. The corresponding values for a total length of 73.7 μ m in experiment 2 were 13.4 μ m (CS), 46.5 μ m (OM) and 13.8 μ m (CS/OM). If the gold particles were uniformly distributed, the expected distributions in CS, OM and CS/OM were calculated from the total number of gold particles counted and the relative length of each segment. The value for *P*, as calculated from χ^2 test [$\chi^2 = (O - E)^2/E$], indicates highly significant labelling of contact sites by both anti-IT and anti-p32 antibodies.

increase the solubility of the signal peptide under the assay conditions (see Fig. 2 methods), it was conjugated to bovine serum albumin (signal-BSA). Signal-BSA inhibited import of precursor into mitochondria (data not shown), indicating that crosslinking of the signal peptide to BSA did not inactivate signal peptide function. When increasing concentrations of signal-BSA were present in the immunoblot reaction there was decreasing reactivity of anti-IT antibodies with p65/63 and p32 (Fig. 2*b*, and quantitation of data in Fig. 2*c*), suggesting that these proteins contain a signal sequence-binding site. Consistent with their alkali extractability, the p65/63 proteins are candidates for signal recognition factor(s) that remain bound to cognate receptor(s) on the mitochondrial surface. This remains to be investigated. By virtue of its resistance to alkali extraction, p32 is a candidate for a membrane-integrated import receptor. We have focused on the characterization of p32 to obtain further anti-IT antibody-independent evidence for its function as a membrane-integrated import receptor. To this end, we purified the anti-IT antibody-reactive p32, raised antibodies against it in rabbits and tested their Fab fragments for the ability to inhibit protein import into mitochondria.

Purification of p32

The starting material for the purification of anti-IT antibody-reactive p32 was alkali-extracted mitochondrial membranes (Fig. 3*Aa*, Load). SDS-solubilized mitochondrial membrane proteins were chromatographed on hydroxylapatite and fractions analysed by SDS-PAGE and either staining with Coomassie blue (Fig. 3*Aa*) or blotting onto nitrocellulose and subsequent probing with anti-IT antibodies (Fig. 3*Ab*) or antibodies against porin (a 30K integral protein of the mitochondrial

outer membrane) (Fig. 3*Ac*). This procedure yielded a useful partial separation of p32 from other mitochondrial membrane proteins.

In a second step, fractions containing the anti-IT antibody-reactive p32 were subjected to reverse-phase HPLC. Subsequent analysis of fractions by SDS-PAGE and either staining with Coomassie blue (Fig. 3*Ba*) or blotting and probing with anti-IT antibodies (Fig. 3*Bb*) revealed that the anti-IT antibody-reactive p32 had been successfully separated from several proteins of similar *M_r*. The anti-IT-reactive p32 was excised from the gel and used to raise antibodies in rabbits. The antibodies thus obtained were monospecific as they reacted with only one protein of 32K (p32) when whole yeast mitochondria were analysed by SDS-PAGE and immunoblotting (Fig. 3*Bd*, lane 2). The corresponding preimmune serum did not react with any of the mitochondrial proteins (Fig. 3*Bd*, lane 1). To investigate whether the anti-p32 antibodies (anti-p32) were directed against the same 32K protein that reacts with the anti-IT antibodies, we tested blots of fractions from the reverse-phase HPLC column. Indeed, both anti-p32 antibodies (Fig. 3*Bc*) and anti-IT antibodies (Fig. 3*Bb*) reacted with a 32K protein with the same elution behaviour, strongly suggesting that the anti-IT and anti-p32 antibodies react with the same protein.

Monospecific anti-p32 inhibit import

The monospecific anti-p32 were able to aggregate mitochondria (data not shown), suggesting that p32 contained epitopes exposed on the outer mitochondrial surface. Fab fragments prepared from purified IgG of anti-p32 were tested in an import reaction containing two different preproteins (Fig. 4*a, b*). In both cases, increasing concentrations of Fab fragments caused an increasing inhibition of import as evidenced by the decreasing amount of mature proteins (lanes 3–7). However, maximum inhibition of import by Fab fragments of anti-p32 was ~75% in both cases tested (Fig. 4*a, b*), unlike the case with Fab fragments of anti-IT antibodies, where inhibition was greater than 95% (Fig. 1) at the highest concentrations included in the assay. The incomplete inhibition of import by anti-p32 Fab might be due to the presence of additional mitochondrial import receptor(s) that were not affected by anti-p32 (ref. 14) but that were affected by anti-IT Fab. As the anti-IT reacted with three proteins on immunoblots, it is possible that 95% inhibition of import by anti-IT Fabs was due to reaction with all three proteins. This would also be consistent with the observation¹⁴ that *mir1*[−] mutant mitochondria can still import proteins, albeit at a reduced rate, even though they do not contain p32. Equivalent concentrations of Fab fragments of IgG from the corresponding preimmune serum had no significant effect on import (Fig. 4*a, b*, lanes 8 and 9). The inhibition of import by the anti-p32 Fab fragments thus further substantiates the identification of p32 as an import receptor.

It should be noted that increasing concentrations of Fab fragments of anti-p32 antibodies resulted in proportionally decreasing quantities of precursors cosedimenting with mitochondria, with an approximately 20% reduction in precursor cosedimentation at the highest concentration of Fab fragments used (Fig. 4*a, b*, compare lanes 2 and 7). A similar reduction in precursor cosedimentation has been observed also with Fab fragments of preimmune or unrelated IgGs (see Fig. 1, compare lane 7 with lanes 9 and 11), suggesting that this is a nonspecific effect.

p32-precursor complex

If p32 functions as a membrane-integrated signal sequence receptor it should be possible to detect a physical association of p32 with the precursor. When a Triton X-100 extract of mitochondria was incubated with radiolabelled pCoxIV-DHFR, the latter was immunoprecipitated by anti-p32 antibodies (Fig. 5, lane 3) but not by anti-IT antibodies (data not shown). Immunoprecipitation of the precursor was indirect in that it was

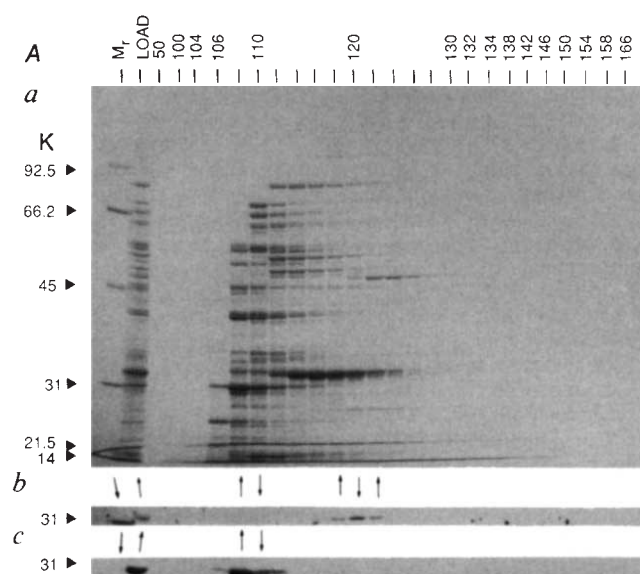
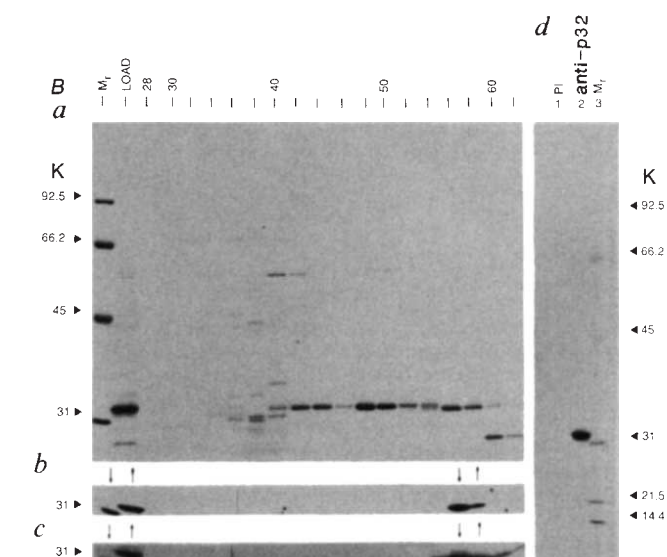


FIG. 3 Purification of p32. **A**, SDS-hydroxylapatite chromatography. Integral membrane proteins of mitochondria were solubilized with SDS and fractionated on an SDS-hydroxylapatite column. Fractions were analysed by SDS-PAGE followed by Coomassie blue staining (**a**) or by western blots probed either with anti-IT antibodies (**b**) or with anti-porin antibodies (**c**). M_r , Unlabelled (**a**) or ^{14}C -labelled (**b** and **c**) M_r weight standards. Numbers along top indicate eluate fractions. Arrows between the panels indicate the corresponding lanes. **B**, Reverse-phase HPLC. Fractions 118–123 of the hydroxylapatite column containing anti-IT-reactive p32 (**Aa** and **Ab**) were subjected to reverse-phase HPLC. Fractions were analysed by SDS-PAGE followed by Coomassie blue staining (**a**) or by western blots probed either with anti-IT antibodies (**b**), or with antibodies against p32 (**c**). M_r , unlabelled (**a**) and ^{14}C -labelled (**b** and **c**) molecular weight standards. **d**, Whole mitochondria analysed by SDS-PAGE and western blotting either with the preimmune IgG (lane 1) or antibodies against p32 (lane 2). Lane 3, ^{14}C -labelled standards. **METHODS**. **A**, Integral membrane proteins obtained after treatment of mitochondria (45 mg) with Na_2CO_3 (see Fig. 2 methods) were solubilized in 4% SDS and adjusted to 50 ml of buffer B (20 mM sodium phosphate, pH 6.8, 0.1% SDS, 1 mM DTT, 0.2 mM PMSF) containing 0.5 mM EDTA and loaded onto a hydroxylapatite²¹ (Bio-Rad) column (1.5 cm $\times$ 18 cm) that was pre-equilibrated with buffer B. The flow rate was adjusted to 18 ml h^{-1} and 1.5-ml fractions were collected. After loading, the column was sequentially washed with 20 ml of buffer B and 50 ml of 0.2 M sodium phosphate, pH



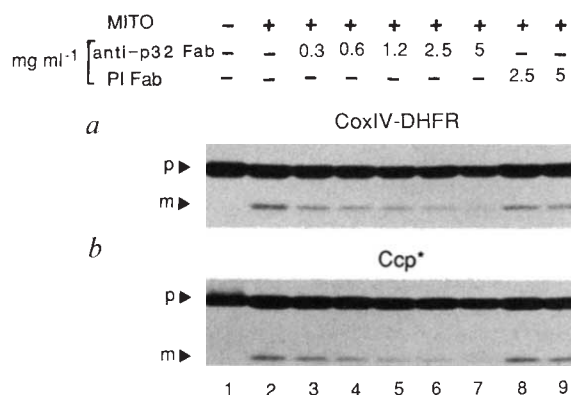
6.8, 0.1% SDS, 1 mM DTT, 0.2 mM PMSF. Elution was by a linear phosphate gradient, 60 ml each of 0.2 M and 0.8 M sodium phosphate, pH 6.8 both containing 0.1% SDS, 1 mM DTT and 0.2 mM PMSF. Fractions were then analysed by SDS-PAGE. The gel was stained with Coomassie blue or blotted onto nitrocellulose and probed either with anti-IT IgG (10 mg ml^{-1} , 1:500) or anti-porin IgG²² (15 mg ml^{-1} , 1:5,000) by procedures described previously¹¹. **B**, SDS-hydroxylapatite column fractions 118–123 containing anti-IT-reactive p32 were pooled and injected onto a C-4 reverse-phase column (Aquapore Butyl, 100 $\times$ 2.1 mm, 7 micron, Applied Biosystems) pre-equilibrated in 60% formic acid. After a 5-min wash with 60% formic acid, the proteins were eluted^{23,24} with a linear gradient of 0–33% acetonitrile in 60% formic acid for a period of 65 min. The gradient from 0 to 13.2% acetonitrile was done in 5 min and continued from 13.2% to 33% acetonitrile over a period of 60 min. Finally, the column was washed with 33% acetonitrile in 60% formic acid for 10 min. The flow rate for the entire period was 0.2 ml min^{-1} , and 0.2 ml fractions were collected from the beginning of the 60% formic acid wash. Two parallel gels containing identical HPLC fractions were run. One (**a**) was analysed by Coomassie blue staining; the other (**b**) was blotted and probed with anti-IT IgG (10 mg ml^{-1} , 1:500) as described previously¹¹. The purified 32K region was cut out of the gel and used for raising anti-p32 antibodies (anti-p32) in rabbits. The specificity of anti-p32 IgG (10 mg ml^{-1} ; 1:4,000) was checked by immunoblotting.

dependent on mitochondrial Triton X-100 extract (lane 2), and specific in that it could be competed by increasing concentrations of signal peptide coupled to bovine serum albumin (signal-BSA) (lanes 4 and 5) but not by corresponding concentrations of uncoupled BSA (lanes 6 and 7). Like signal-BSA, anti-IT Fab

FIG. 4 Monovalent Fab fragments of anti-p32 IgG inhibit import. Post-ribosomal supernatant obtained from a reticulocyte lysate translation mixture containing a precursor, either pCoxIV-DHFR (**a**) or modified pre-cytochrome *c* peroxidase (pCcp*, **b**), was incubated in a total volume of 0.15 ml at 30 $^{\circ}\text{C}$ with isolated yeast mitochondria (50 μg protein) preincubated in the absence (lane 2) or presence of various amounts (45–750 μg) of Fab fragments isolated either from anti-p32 IgG (lanes 3–7) or preimmune IgG (lanes 8 and 9). Lane 1 shows 25% of the PRS used in the import assay. p, Precursor; m, mature form. Numbers in the top panel indicate final concentrations of Fab (mg ml^{-1}).

METHODS. Plasmid pGT-CoxIV-DHFR is described in Fig. 1 methods. Plasmid pSP6CCP containing the entire coding sequence for pre-cytochrome *c* peroxidase² (pCcp) under the SP6 promoter was provided by Dr J. Kaput. The presumed haem-binding site of pCcp was removed by digesting the plasmid pSP6CCP with *Nco*I and religating it. The resulting plasmid pSP6CCP* encodes a modified cytochrome *c* peroxidase precursor (pCcp*) which lacks amino acids 126 (Trp) through to 190 (Pro) of the mature cytochrome *c* peroxidase². The plasmid pSP6CCP* was linearized downstream from the gene with *Eco*RI and transcribed using SP6 RNA polymerase as described⁵. ^{35}S -labelled pCoxIV-DHFR and pCcp* were synthesized *in vitro* in a reticulocyte lysate translation system and import was performed (as described in Fig. 1 methods) in the absence or presence of Fab fragments

was also able to inhibit precursor coprecipitation by anti-p32 (data not shown). These data are consistent with p32 being a signal sequence receptor. When the Triton X-100 extract of mitochondria was centrifuged at 365,000g for 1 h, about 90% of p32 remained in the supernatant fraction (data not shown).



isolated from IgG of either decomplexed preimmune serum or decomplexed immune serum against purified p32. Radioactivity in p and m in dried gels was measured with an AMBIS β -scanner.

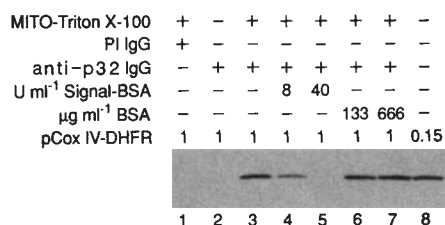


FIG. 5 Immunoprecipitation of a p32-precursor complex with anti-p32 antibodies. Post-ribosomal supernatant (PRS) containing pCoxIV-DHFR was incubated in the absence or presence of a Triton X-100 extract of mitochondria (MITO-Triton X-100) and in the absence or presence of indicated quantities of either BSA or signal-BSA. After incubation, mixtures were immunoprecipitated either with anti-p32 IgG or IgGs derived from preimmune serum (PI). Lane 8 shows 15% of the PRS used in the immunoprecipitation reactions. METHODS. Mitochondria (1 mg protein) were solubilized on ice for 15 min in 0.3 ml buffer containing 20 mM HEPES/KOH, pH 7.5, 15 mM EDTA, 0.15 M KCl, 2 mM DTT, 10 units ml⁻¹ trasylol, 12.5 μg ml⁻¹ each of antipain,

chymostatin, leupeptin and pepstatin, 1 mM PMSF and 1% Triton X-100. The mixture was centrifuged in a microcentrifuge (14,000g) for 20 min at 4 °C and the supernatant represented the Triton X-100 extract of mitochondria. A 4.5 μl aliquot of this extract was incubated at 25 °C for 30 min in a total volume of 150 μl in a binding buffer that contained, at final concentration, 20 mM HEPES/KOH, pH 7.5, 5 mM EDTA, 4.5 mM KCl, 45.5 mM CH₃COOK, 60 μM DTT, 0.1% Triton X-100, 10 μM ATP, 10 units ml⁻¹ trasylol, 5 μg ml⁻¹ each of antipain, chymostatin, leupeptin and pepstatin, 0.5 mM PMSF and 1 μl PRS containing ³⁵S-labelled pCoxIV-DHFR and, where indicated, the appropriate amount of either signal-BSA (8 or 40 units ml⁻¹) or corresponding concentrations of uncoupled BSA (133 or 666 μg ml⁻¹). Thereafter, 500 μg of preimmune IgG or anti-p32 IgG in binding buffer, lacking PRS, DTT and ATP, was added to each reaction to a final volume of 210 μl. After incubation at 4 °C for 1 h, samples were centrifuged at 4 °C for 2 min at 14,000g to remove aggregates. Supernatants were removed and incubated for 90 min at room temperature with 10 mg protein A-Sepharose (Pharmacia) in binding buffer, lacking PRS, DTT and ATP, to yield a final volume of 600 μl. The protein A-Sepharose beads, with bound antigen-antibody complex, were washed four times with PRS and ATP-free binding buffer containing 15 μM DTT and twice with 20 mM HEPES/NaOH, pH 8.0, 50 mM NaCl, 5 mM EDTA, 15 μM DTT. The samples were analysed by 12% SDS-PAGE and fluorography.

Thus, p32 in the detergent extract was not associated with large submitochondrial particles. We cannot, however, rule out the possibility that after detergent solubilization of mitochondria, p32 remains complexed to other protein(s) (such as p65/63) that also contain signal sequence-binding sites, resulting in coprecipitation of precursor by anti-p32.

Immunoelectron microscopy

Evidence that translocation occurs at contact sites has previously been provided by immunoelectron microscopic localization of a translocation intermediate to these sites⁴. We performed immunogold electron microscopy (Fig. 6) using isolated yeast mitochondria incubated with IgG of either anti-IT, anti-p32, or corresponding preimmune sera or an antiserum against a crude outer mitochondrial membrane fraction (anti-OM). The preimmune sera did not yield immunogold decoration of mitochondria (data not shown). With both anti-IT and anti-p32 there was decoration of only a few sites on the mitochondrial surface

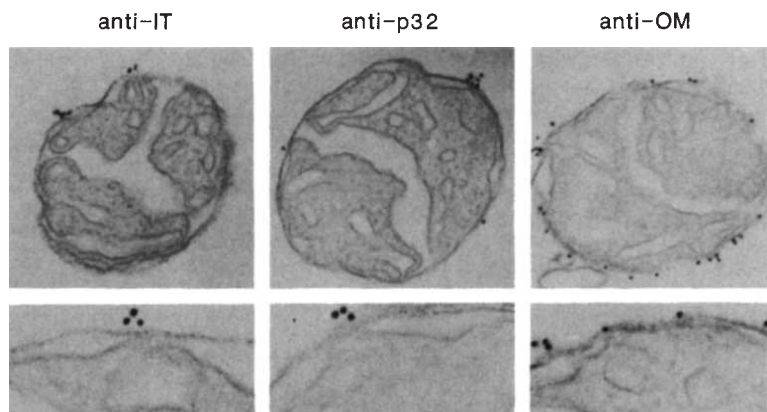
which coincided with contact sites between the outer and inner mitochondrial membranes. However, anti-OM decorated the entire mitochondrial surface. Morphometric analysis of the distribution of gold particles on the mitochondrial surface indicated that the incidence of labelling at contact sites by both anti-IT and anti-p32 antibodies was highly significant (Table 1).

Subfractionation of mitochondria

To obtain additional evidence for p32 being located in contact sites, we disrupted mitochondria, separated membrane subfractions by isopycnic sucrose gradient centrifugation and analysed the resulting gradient fractions by SDS-PAGE and immunoblotting with antibodies against porin (an outer membrane protein), cytochrome *c* oxidase subunit II (an inner membrane marker) as well as p32. It has been shown that a membrane subfraction containing both inner and outer mitochondrial membrane linked together via contact sites bands at a slightly lower density than the inner membrane fraction^{4,15}. The data in Fig. 7 demonstrate

FIG. 6 Immunogold electron microscopy of isolated yeast mitochondria. Mitochondria were immobilized, fixed with paraformaldehyde and incubated with either anti-IT antibodies, monospecific antibodies against the p32 (anti-p32) or polyspecific antibodies against SDS-denatured proteins of a crude mitochondrial outer membrane fraction (anti-OM), and then with goat anti-rabbit IgG conjugated to colloidal gold. Magnifications: upper panels ×35,000; lower panels, ×80,000.

METHODS. Isolated mitochondria (500 μg protein) in buffer C (20 mM HEPES/KOH, pH 7.5, 0.6 M sorbitol, 1 mM EDTA) containing 0.5 mM PMSF were allowed to settle onto a poly-L-lysine-coated polystyrene dish for 1 h at 4 °C. Unsettled mitochondria were decanted and the plate was immersed in ice-cold 2% paraformaldehyde in buffer C for 30 min at 4 °C. Fixed mitochondria were successively washed with buffer C containing 10 mM NH₄Cl and buffer C containing 1% BSA. The samples were then incubated either with anti-IT IgG, anti-p32 IgG, anti-OM IgG or preimmune IgG, each at a dilution of 1:100 in buffer C containing 1% BSA. Anti-OM antibodies were raised against SDS-denatured proteins of a crude mitochondrial outer membrane fraction²⁵. All IgG fractions were made from decomplexed antisera and the protein concentrations were adjusted to 10 mg ml⁻¹ before use. Following incubation for 2 h at 4 °C, the samples were washed with buffer C containing 1% BSA and incubated with goat IgG (10 μg ml⁻¹) in the same buffer for 10 min. The samples were rewashed with buffer C containing 1% BSA, and incubated for 90 min at 4 °C with goat anti-rabbit IgG conjugated to colloidal gold (10 nm, 1:25, Janssen). The samples were washed with buffer C, immersed in Karnovsky's aldehyde fixative containing



0.6 M sorbitol (12 h at 4 °C), osmicated (1 h at 4 °C), stained with uranyl acetate (1 h at room temperature) and dehydrated in a graded series of ethanol washes as described²⁶. The samples on the plastic dishes were scored with a scalpel, flushed off with cold propylene oxide, and pelleted by centrifugation. The pellets were immersed in an EPON:propylene oxide mixture (1:1), infiltrated with pure EPON and embedded. Ultra-thin sections (70–90 nm) were mounted on copper grids and stained with uranyl acetate. The samples were examined with a JOEL 100CX electron microscope at an accelerating voltage of 80 kV.

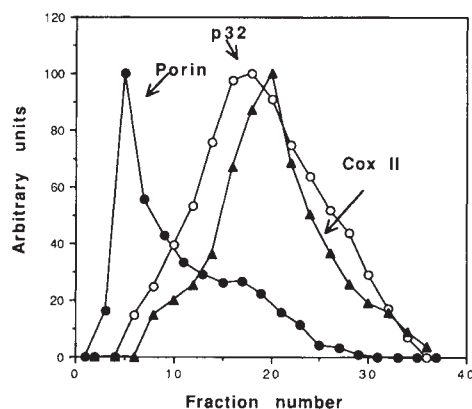


FIG. 7 Separation of submitochondrial fractions by isopycnic sucrose gradient centrifugation. Membrane subfractions from disrupted mitochondria were separated on a sucrose gradient, and subjected to SDS-PAGE, blotted onto nitrocellulose and probed with either anti-porin (●), anti-p32 (○) or anti-cytochrome *c* oxidase subunit II (▲). The immunoblots were quantitated by densitometric scanning. Values obtained were converted into arbitrary units and plotted against the corresponding fraction number. The peak for each protein was considered to be 100 arbitrary units.

that p32 indeed peaked in fractions of lower density than the cytochrome *c* oxidase subunit II of the inner membrane, consistent with the immunoelectron microscopic localization of p32 to contact sites.

Discussion

We have prepared anti-IT antibodies to mimic the signal peptide of a mitochondrial precursor protein. Fab fragments of these antibodies inhibited protein import into mitochondria. On immunoblots, the anti-IT antibodies reacted with three mitochondrial proteins (p65, p63 and p32). Two of these proteins (p65 and p63) are extractable at pH 11.5 and remain to be further characterized. The other anti-IT-reactive protein, p32, is an integral membrane protein not extractable by alkali.

The anti-IT-reactive p32 was purified to homogeneity and monospecific antibodies were raised against purified (and SDS denatured) p32. Fab fragments prepared from anti-p32 IgG inhibited import. Moreover, radiolabelled precursor, when incubated with a detergent extract of mitochondria containing p32, could be specifically immunoprecipitated by anti-p32 consistent with p32 being a signal sequence receptor.

Aggregation of mitochondria by anti-p32 and inhibition of import by anti-p32 Fab suggested an outer membrane location for p32. Immunoelectron microscopy using anti-IT or anti-p32 yielded decoration of outer membrane restricted to contact sites

METHODS. Submitochondrial fractionation was done essentially as described¹⁵. Briefly, mitochondria (1 ml, 23 mg ml⁻¹) were diluted with 24 ml of 20 mM HEPES/KOH, pH 7.5 containing 0.5 mM EDTA, 1 µg ml⁻¹ each of antipain, chymostatin, leupeptin and pepstatin and 1 mM PMSF, and swollen by incubation for 30 min on ice. Sucrose solution (1.8 M) in the above buffer was then added to a final concentration of 0.45 M, incubated on ice for 10 min and sonicated in aliquots (6–7 ml) in a cell disrupter equipped with a microtip for 2 min on ice at 80% duty cycle. Sonicated material was centrifuged at 32,000g for 20 min at 4 °C to remove unbroken mitochondria and large aggregates. The supernatant was removed, and spun again in a Beckman Ti 50.2 rotor at 200,000g for 1 h. The pellet containing the submitochondrial membrane fraction was resuspended in 5 mM HEPES/KOH, pH 7.5, 10 mM KCl and 0.1 mM PMSF (1–1.2 ml) using a Dounce homogenizer. Resuspended membranes were loaded onto a linear sucrose gradient (35.2 ml, 0.85–1.6 M sucrose in 5 mM HEPES/KOH, pH 7.5, 10 mM KCl, 0.1 mM PMSF), and centrifuged at 4 °C in a SW 28 rotor at 96,000g for 20 h. Fractions (1 ml) were collected from the top. Aliquots (100 µl) were diluted with water to 540 µl, proteins were precipitated with 10% trichloroacetic acid and processed for SDS-PAGE. Following electrophoresis on 12% polyacrylamide gels, samples were blotted onto nitrocellulose and probed with either anti-porin IgG²² (15 mg ml⁻¹, 1:5,000), anti-p32 IgG (10 mg ml⁻¹, 1:4,000) or antiserum against subunit II of cytochrome *c* oxidase (1:2,500). The immunoblot procedure was as described¹¹. The immunoblots were quantitated by densitometric scanning as described in Fig. 2 (c) methods. The obtained values were converted into arbitrary units and plotted against the corresponding gradient fraction number.

where the outer membrane is closely apposed to the inner membrane. Location of p32 at membrane contact sites was also suggested by isopycnic centrifugation of disrupted mitochondrial membranes. Several proteins have so far been identified as candidates for components of the import machinery. A yeast mitochondrial outer membrane protein of 42K was shown to be photo-crosslinked to a chimaeric precursor protein⁷. Fab fragments of antibodies against a *Neurospora crassa* 19K mitochondrial outer membrane protein were found to inhibit import of most but not all mitochondrial precursors⁸. Neither the 42K nor the 19K protein was found to be enriched in contact sites. Chemical crosslinking studies with a synthetic mitochondrial signal peptide have resulted in the identification of a 30K integral protein in the rat heart mitochondrial outer membrane as a putative signal sequence receptor⁹. Recently, a 29K protein of rat liver mitochondrial membranes has been partially purified by affinity chromatography using a synthetic signal peptide and proposed to be a putative import receptor¹⁰. The relationship between p32 and the 30K protein of rat heart mitochondria or the 29K protein of rat liver mitochondria remains to be investigated. We envisage that p32 with its signal sequence-binding domain is a subunit of a protein-conducting channel in the outer membrane that can be linked up with a protein-conducting channel in the inner membrane permitting the conductance of proteins from cytoplasm to mitochondrial matrix¹⁶. □

Received 18 May; accepted 14 August 1990.

- Maccacchini, M. L., Rudin, Y., Blobel, G. & Schatz, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 343–347 (1979).
- Kaput, J., Goltz, S. & Blobel, G. *J. biol. Chem.* **257**, 15054–15058 (1982).
- Viebrock, A., Perz, A. & Sebald, W. *EMBO J.* **1**, 565–571 (1982).
- Schwaiger, M., Herzog, V. & Neupert, W. *J. Cell Biol.* **105**, 235–246 (1987).
- Murakami, H., Pain, D. & Blobel, G. *J. Cell Biol.* **107**, 2051–2057 (1988).
- Ohba, M. & Schatz, G. *EMBO J.* **6**, 2109–2115 (1987).
- Vestweber, D., Brunner, J., Baker, A. & Schatz, G. *Nature* **341**, 205–209 (1989).
- Söllner, T., Griffiths, G., Pfanner, R., Pfanner, N. & Neupert, W. *Cell* **59**, 1061–1070 (1989).
- Gillespie, L. L. *J. biol. Chem.* **262**, 7939–7942 (1987).
- Ono, H. & Tuboi, S. *J. Biochem.* **107**, 840–845 (1990).
- Pain, D., Kanwar, Y. S. & Blobel, G. *Nature* **331**, 232–237 (1988).
- Maarse, A. C. *et al.* *EMBO J.* **3**, 2831–2837 (1984).
- Hurt, E. C., Pesold-Hurt, B. & Schatz, G. *FEBS Lett.* **178**, 306–310 (1984).
- Murakami, H., Blobel, G. & Pain, D. *Nature* **347**, 488–491 (1990).
- Pon, L., Moll, T., Vestweber, D., Marshallsay, B. & Schatz, G. *J. Cell Biol.* **109**, 2603–2616 (1989).
- Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1496–1500 (1980).
- Pain, D. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3288–3292 (1987).

- Pelham, H. R. B. & Jackson, R. J. *Eur. J. Biochem.* **67**, 247–256 (1976).
- Kaput, J., Brandriss, M. C. & Prussak-Wieckowska, T. *J. Cell Biol.* **109**, 101–112 (1989).
- Posnett, D. N., McGrath, H. & Tam, J. P. *J. biol. Chem.* **263**, 1719–1725 (1988).
- Berrios, M., Filson, A. J., Blobel, G. & Fisher, P. A. *J. biol. Chem.* **258**, 13384–13390 (1983).
- Mihara, K., Blobel, G. & Sato, R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7102–7106 (1982).
- Shelness, G. S., Kanwar, Y. S. & Blobel, G. *J. biol. Chem.* **263**, 17063–17070 (1988).
- Heukeshoven, J. & Dernick, R. *J. Chromatogr.* **326**, 91–101 (1985).
- Riezman, H. *et al.* *EMBO J.* **2**, 1105–1111 (1983).
- Kanwar, Y. S. & Farquhar, M. G. *J. Cell Biol.* **81**, 137–153 (1979).

ACKNOWLEDGEMENTS. We thank S. deSanctis for synthesizing the peptide, G. S. Shelness for help with reverse-phase HPLC and H. Shio for help with electron microscopy. We acknowledge A. P. G. M. Van Loon for the pGT-CoxIV-DHFR plasmid, J. Kaput for the pSP6CCP plasmid and for providing information necessary for making the pSP6CCP* construct, and K. Mihara for antibodies against porin and subunit II of cytochrome *c* oxidase. We thank W. Tam for technical help, D. Schnell, N. Schülke, D. Schwartz, N. Chaudhary and C. Nicchitta for helpful comments and critical reading of the manuscript.

The distribution of iron in the Perseus cluster

T. J. Ponman, D. Bertram, M. J. Church, C. J. Eyles,
M. P. Watt, G. K. Skinner & A. P. Willmore

School of Physics and Space Research, University of Birmingham,
Birmingham B15 2TT, UK

THE hot ($\sim 10^8$ K) gas that permeates clusters of galaxies may be primordial matter left over from the initial density perturbation that grew into the cluster, or may be gas ejected from cluster galaxies. The detection of iron emission features in cluster X-ray spectra proves that at least some of the intracluster medium (ICM) is not primordial, and much theoretical modelling of clusters has been based on the assumption that the entire ICM derives from the galaxies. We report here hard-X-ray spectral imaging observations of the Perseus cluster obtained with the coded-mask X-ray telescope on Spacelab 2 which reveal an iron abundance gradient in the ICM. Most of the ICM seems to be primordial, with the iron-rich gas confined largely to the cluster core where it was probably deposited by stripping of cluster galaxies. The total mass of iron required to explain the X-ray spectra is substantially smaller than has been assumed on the basis of a chemically homogeneous ICM.

The coarse component of the dual X-ray telescope on Spacelab 2 (ref. 1) used coded-mask imaging² to achieve angular resolution of $12'$ over a 6° field together with proportional-counter energy resolution covering the 2–32 keV band, recording $\sim 7 \times 10^5$ photons from the Perseus cluster during an observation in July–August 1985. Here these data have been reduced to a spectral image with $3'$ pixels and 42 energy channels spanning the range 3.5–25 keV.

The spectrum of the integrated flux out to $42'$ from the cluster centre shows emission at ~ 6.7 keV in excess of that for a simple bremsstrahlung continuum, arising from the K-line blend of ionized iron. Fitting this spectrum with models of thermal emission from a hot optically thin plasma³, allowing for the effects of the cluster redshift of $z = 0.0182$, we derive a temperature of 6.2 ± 0.4 keV (1σ errors) and an iron abundance, in solar units, of $Z = 0.38 \pm 0.12$ (equivalent width, EW = 485 eV), in good agreement with previous results^{4,5}.

To investigate the radial variation of the abundance, the spectral image was binned into $6'$ annuli centred on NGC1275, which lies at the cluster centre. NGC1275 is an active galaxy known to be a power-law point source of X-rays^{6,7}. Given the $16'$ full width to zero response of our point-spread function, the contribution from this galaxy will be confined largely to the innermost bin, where we have therefore included a power-law component in the spectral model. The strength of this point-source component is constrained to better than 20% by the three-dimensional fitting procedure described below and lies in the range observed by previous instruments^{7–9}. Spectra from each annulus and best-fit models are shown in Fig. 1. The strength of the iron emission decreases with increasing radius, and the implied metal abundances are shown in Fig. 2a.

Although there have been previous indications that an iron abundance gradient might be present in the Perseus cluster from observations obtained by the slit-scanning instrument Spartan 1 (ref. 9), our results represent the first unambiguous detection of such a gradient in a galaxy cluster. Initial analysis of the Spartan data gave an abundance of $0.81^{+0.26}_{-0.16}$ for events recorded with the slit centre within $5'$ of the cluster centre, and $0.41^{+0.41}_{-0.25}$ for the region $r = 6\text{--}20'$ (90% confidence errors).

The simple radial binning procedure will tend to underestimate gradients in the cluster gas, because these will be integrated over a range of radii along the line of sight, and also blurred by the $\sim 12'$ resolution of the telescope. To allow for this, a three-dimensional model of the cluster gas was constructed (see

ref. 10 for details). The models consist of three main components: the hot ICM, a central point source, and a cooling flow surrounding it. The emergent emission is projected and folded through the spatial and spectral response of the instrument before comparison with the observed spectral image, and the model parameters optimized using standard χ^2 fitting techniques¹¹.

The ICM gas density ρ is parameterized using a spherically symmetric core-index model $\rho(r) = \rho_0(1 + (r/a)^2)^{-\alpha_\rho}$ (where r is the radius in arcminutes, ρ_0 the central density, α_ρ the density index, and the core radius a was fixed at 9.1 arcmin from the Einstein results¹²) and the temperature represented as a power-law function of radius $T(r) = T_1 r^{-\alpha_T}$ where T_1 is the temperature at a radius corresponding to $1'$ and α_T is the temperature index. Emission from NGC1275 was included in our model by fitting a central point-source component with a power-law photon spectrum $N(E) = N_1 E^{-\alpha_{ps}}$ (where N_1 is a normalization constant and α_{ps} is the photon index) and an intrinsic absorbing column¹³ of $4.3 \times 10^{21} \text{ cm}^{-2}$. NGC1275 is surrounded by a strong cooling flow^{6,8} which we have insufficient spatial resolution to map. We therefore adopted the parameters derived from the Einstein observations. The most reliable density profile within the cooling flow (A. C. Fabian, personal communication) is that derived from the Imaging Proportional Counter (IPC) (ref. 14), whereas for the temperature profile in the flow we have adopted $T \propto r^{0.5}$, as suggested by deprojection of the High Resolution Imager data⁶. This profile is normalized by requiring it to be continuous with our ICM temperature profile at the outer radius r_c of the cooling flow. Unfortunately r_c is uncertain, lying in

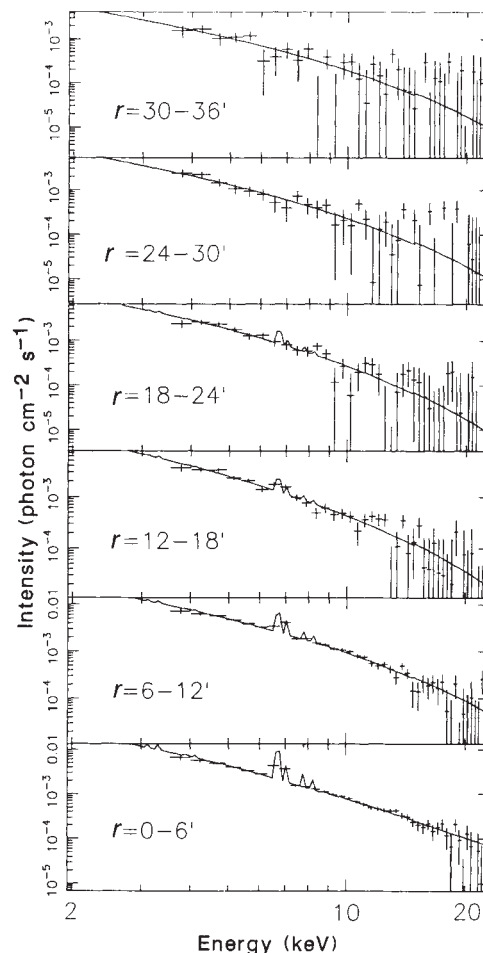


FIG. 1. Spectra from successive $6'$ annular elements together with best-fit hot plasma models. The line strength decreases steadily with increasing radius.

the range⁶ 5–10'. The IPC density matches our fitted profiles for the ICM at a radius of $\sim 5'$, which we have therefore adopted for r_c . Extending the IPC profile out to 10' has less of an effect on our results than some of the other uncertainties discussed below.

The metallicity of the gas (in both the cooling flow and the intracluster medium) was parameterized using various simple analytical forms (such as gaussian, top hat, ramp). These models give similar results for the total amount of iron and its degree of concentration toward the cluster centre, and our limited spatial resolution does not permit us to discriminate between them. We adopt a top-hat function, $Z = Z_0$ for $r < R_Z$, $Z = 0$ for $r > R_Z$, as our standard model, where R_Z must be thought of as the characteristic radius of the metallicity distribution (an abrupt cutoff is not required by our data).

The above model gives an acceptable fit ($\chi^2 = 23,427$ with 23,213 degrees of freedom) to our spectral image out to 40' from NGC1275, where the signal becomes too weak to constrain the spectral properties. A uniform distribution of iron is overwhelmingly ($P \gg 99.9\%$) rejected, and the best-fit parameters with 1σ single-parameter errors¹¹ are $\rho_0 = 4.1^{+0.2}_{-0.1} \times 10^{-3} \text{ cm}^{-3}$, $\alpha_p = 0.88 \pm 0.02$, $T_1 = 10.9^{+1.5}_{-1.1} \text{ keV}$, $\alpha_T = 0.17 \pm 0.05$, $N_1 = 4.1^{+0.6}_{-0.5} \times 10^{-2} \text{ photon cm}^{-2} \text{ s}^{-1} \text{ keV}^{-1}$, $\alpha_{ps} = 2.1 \pm 0.2$, $Z_0 = 2.09^{+0.07}_{-0.12}$, $R_Z = 9.3^{+0.5}_{-0.6}$. (A distance to the cluster of 109 Mpc and an interstellar column (A. A. Stark, C. Heiles, J. Bally and R. Linke, manuscript in preparation) of $1.4 \times 10^{21} \text{ cm}^{-2}$ were assumed throughout.)

Although the formal statistical errors on the two metallicity parameters are small, possible errors in the assumed properties of the cooling flow introduce larger uncertainties. Z_0 is sensitive to the density in the cooling flow (an approximately inverse-square dependence) which fortunately should be quite well determined by the Einstein data within 5' of the centre. The temperature of the flow has a much weaker effect (power-law profiles between $r^{0.3}$ and $r^{0.7}$ were investigated). In general, credible variations in the cooling flow properties lead to changes of up to 15% in R_Z and 25% in Z_0 , although the possibility that the flow might have a range of different densities and temperatures at each radius¹⁵ (a 'multiphase' flow) could have still larger effects on Z_0 . Systematic errors arising from uncertainties in instrument calibration and residual background are negligible compared to these effects.

The high central metallicities apparent from the Spacelab 2 and Spartan data disagree with the results obtained by the Einstein Solid State Spectrometer¹⁶, which observed the core of the cluster (with a 3' radius aperture) at lower energies and gave an iron abundance estimate of 0.44 ± 0.20 in solar units. The Einstein result is, however, vulnerable to uncertainties in the structure of the cooling flow (because the low-energy line emission originates from the cooler components of the flow, whereas much of the continuum arises from the hotter phases), and moreover would be substantially increased by inclusion of a power-law point-source component at a level similar to that commonly observed (R. F. Mushotzky, personal communication). Einstein Focal Plane Crystal Spectrometer data cannot constrain the iron abundance, but suggest that oxygen is more abundant than iron¹⁷. Our results then imply a high central oxygen abundance.

Figure 2a suggests a complete absence of iron in the intracluster medium at large radii. To check this, we added a uniform background iron abundance to our model (that is, $Z = Z_0$ for $r < R_Z$ and $Z = Z_b$ for $r > R_Z$). The best fit was obtained with $Z_b = 0$, and we can set an upper limit of $Z_b = 0.10$ (90% confidence) which is insensitive to details of the properties of the cooling flow.

Iron line emission with equivalent widths of $\sim 100\text{--}300 \text{ eV}$ seems to be a fairly common feature of Seyfert galaxies¹⁸. Such emission, if present in NGC1275, could account for only a small fraction of the strong line we observe. Much stronger emission ($\text{EW} \approx 1.5 \text{ keV}$) has been observed in a few edge-on Seyfert 2 galaxies¹⁹. We have therefore tested the effects of adding an

iron line to the continuum of our central point source. Our data allow a substantial line at $\sim 6.4 \text{ keV}$, provided that it contributes no more than 50% of the total iron emission. The remainder must be distributed within a characteristic radius $R_Z \approx 9.5'$ much as above—a model combining an iron point source with uniform metallicity in the cluster gas is not consistent with the data. The main impact of introducing a central iron source is therefore to reduce Z_0 by a factor of up to two.

If iron is concentrated towards the centre of clusters, then (owing to the higher emission measure at higher density) less of it is required to produce the observed X-ray emission lines²⁰. This effect can be seen in Fig. 2b, which shows the cumulative iron mass for the best-fit top-hat and gaussian iron distributions, and also for the case of a uniform abundance of 0.38. The total mass of iron implied by our results, $\sim 3.3 \times 10^{10} h_{50}^{-5/2} M_\odot$ (where h_{50} is the Hubble constant in units of $50 \text{ km s}^{-1} \text{ Mpc}^{-1}$) for our best fits, is much smaller than the result for a chemically homogeneous ICM, particularly when extrapolated to larger radii.

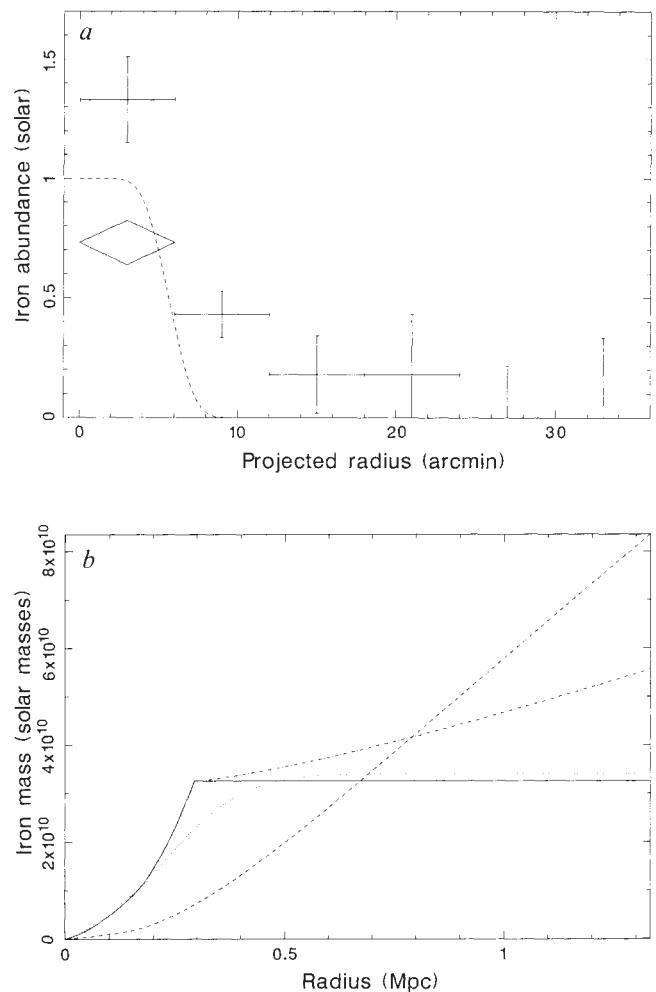


FIG. 2 *a*, Iron abundance as a function of annular radius from the fits displayed in Fig. 1. Error bars and upper limits are at the 1σ level. For the innermost bin, omitting the power-law central source from the model results in a reduced abundance, plotted as a diamond. The point-spread function of the instrument (dashed) is shown centred on $r=0$ for comparison. *b*, Cumulative iron mass as a function of radius in the cluster for the best-fit top-hat (solid) and gaussian (dotted) iron distributions, together with the distribution obtained for a uniform abundance of 0.38 solar (dashed). The top-hat and gaussian models give very similar estimates of the iron mass. Inclusion of our maximum iron background, $Z_b = 0.1$ (dot-dashed) gives a larger iron mass, which at large radii tends towards 1/4 of that for uniform metallicity.

Despite uncertainties in detail resulting from the cooling flow and central source, our results establish that the metallicity of the ICM is high ($>$ solar) in the cluster core, and low (≤ 0.1 solar) outside it. This suggests that the gas in the core has been ejected from galaxies, whereas much of the rest is primordial gas that is left over from inefficient galaxy formation or has fallen into the cluster afterwards. The temperature and density distributions we derive for the gas¹⁰ are in reasonable agreement with the results of hydrodynamic studies of both processes^{21,22} and imply either a rather low efficiency of galaxy formation (we find $M_{\text{gas}}/M_{\text{gal}} > 4$) or a high density of material around the collapsing protocluster.

How has the iron collected in the cluster core? Diffusive sedimentation of the heavier iron nuclei²⁰ occurs too slowly to establish a significant gradient by the present time²³. The radial variation in the relative abundance of spiral and elliptical galaxies could lead to a metallicity gradient²⁴, but would produce an effect far weaker than we observe, given the large quantity of primordial gas that seems to be present. The most likely explanation seems to be stripping of gas from galaxies owing to motion through the higher density medium near the cluster centre.

There is, however, a difficulty with this interpretation. Although gas stripping is a complex and imperfectly understood process, hydrodynamic studies (see ref. 25 for a review) indicate that it should become effective at ambient densities somewhere in the range $n_0 \approx 10^{-3}$ – 10^{-4} cm⁻³, a result supported by the observation²⁶ of what seems to be gas being stripped from M86 in the Virgo cluster at ICM densities of $\sim 5 \times 10^{-4}$ cm⁻³. This density range corresponds to radii of 18'–70' (ref. 10), considerably larger than that within which we find the iron to be con-

tained. It is possible that stripping at larger radii has been suppressed by the presence of massive galaxy halos²⁷ which are disrupted close to the cluster core²⁸. □

Received 5 February; accepted 9 August 1990.

1. Skinner, G. K. *et al. Astr. Lett. Comm.* **27**, 199–209 (1988).
2. Skinner, G. K. *Nucl. Instrum. Meth.* **221**, 33–40 (1984).
3. Raymond, J. C. & Smith, B. W. *Astrophys. J. Suppl.* **35**, 419–439 (1977).
4. Mushotzky, R. F., Serlemitsos, P. J., Smith, B. W., Boldt, E. A. & Holt, S. S. *Astrophys. J.* **225**, 21–39 (1978).
5. Okumura, Y. *et al. Publ. astr. Soc. Japan* **40**, 639–651 (1988).
6. Fabian, A. C., Hu, E. M., Cowie, L. L. & Grindlay, J. *Astrophys. J.* **248**, 47–54 (1981).
7. Primini, F. A. *et al. Astrophys. J.* **243**, L13–L17 (1981).
8. Branduardi-Raymont, G. *et al. Astrophys. J.* **248**, 55–60 (1981).
9. Ulmer, M. P., Cruddace, R. G., Fenimore, E. E., Fritz, G. G. & Snyder, W. A. *Astrophys. J.* **319**, 118–125 (1987).
10. Eyles, C. J. *et al. Astrophys. J.* (submitted).
11. Lampton, M., Margon, B. & Bowyer, S. *Astrophys. J.* **208**, 177–190 (1976).
12. Jones, C. & Forman, W. *Astrophys. J.* **276**, 38–55 (1984).
13. Gorenstein, P. *et al. Astrophys. J.* **224**, 718–723 (1978).
14. Branduardi-Raymont, G. *et al. Adv. Space Res.* **5**, 133–136 (1985).
15. Thomas, P. A., Fabian, A. C. & Nulsen, P. E. *J. Mon. Not. R. astr. Soc.* **228**, 973–991 (1987).
16. Mushotzky, R. F., Holt, S. S., Smith, B. W., Boldt, E. A. & Serlemitsos, P. J. *Astrophys. J.* **244**, L47–L51 (1981).
17. Canizares, C. R., Markert, T. H. & Donahue, M. E. in *Cooling Flows in Clusters and Galaxies* (ed. Fabian, A. C.) 63–72 (Kluwer, Dordrecht, 1988).
18. Pounds, K. A., Nandra, K., Stuart, G. C. & Leighly, K. *Mon. Not. R. astr. Soc.* **240**, 769–783 (1989).
19. Koyama, K. *et al. Publ. Astr. Soc. Japan* **41**, 731–737 (1989).
20. Fabian, A. C. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **181**, 5P–7P (1977).
21. Cowie, L. L. & Perrenod, S. C. *Astrophys. J.* **219**, 354–366 (1978).
22. Perrenod, S. C. *Astrophys. J.* **224**, 285–292 (1978).
23. Rephaeli, Y. *Astrophys. J.* **225**, 335–338 (1978).
24. Pastor, J., Canal, R., Sanahuja, B. & Pello, R. *Astrophys. Space Sci.* **157**, 221–226 (1989).
25. Sarazin, C. L. *Rev. mod. Phys.* **58**, 1–115 (1986).
26. Forman, W., Schwarz, J., Jones, C., Liller, W. & Fabian, A. C. *Astrophys. J.* **234**, L27–L31 (1979).
27. Mathews, W. G. *Astr. J.* **97**, 42–56 (1989).
28. Merritt, D. *Astrophys. J.* **276**, 26–37 (1984).

ACKNOWLEDGEMENTS. We thank NASA for the operation of the Spacelab 2 mission, SERC for financial support for development of the X-ray telescope, and R. Cruddace and R. Mushotzky for unpublished results. Analysis of the data was performed on the Starlink node at Birmingham.

A family of cometary globules around an infrared source near the Rosette nebula

David L. Block

Department of Computational and Applied Mathematics,
University of the Witwatersrand, 1 Jan Smuts Avenue,
2001 Johannesburg, South Africa

SEEN in silhouette against the numerous ionized H II regions in the Milky Way are groups of small dark clouds of neutral gas and dust known as Bok globules. In some cases the stellar wind from a nearby star strips material from the globules, creating a tail pointing away from the source of the wind. Here I report the use of a photographic enhancement technique to reveal a family of four such 'cometary' globules in the southeast quadrant of the Rosette nebula, NGC2237–2246. The globules are not visible on R-glass copies of the National Geographic/Palomar Observatory Sky Survey. The tails of the globules all point away from the IRAS (Infrared Astronomy Satellite) source 06314+0427, the position of which coincides with the peak of CO emission¹. The far-infrared luminosity of 06314+0427 is estimated to be a factor of at least 880 greater than the Sun's luminosity, with a total energy output enough to drive a stellar wind that could produce the observed globules.

Malin^{2,3} and Madsen⁴ have demonstrated the potential of using photographically amplified prints to reveal features of extremely low surface brightness on astronomical plates, which are neither apparent on visual inspection nor detectable by conventional microdensitometry. Because faint images tend to be located in the uppermost layers of the developed emulsion, whereas grains owing to chemical fog are randomly distributed throughout the emulsion coating, contact copies of original plates can be made, recording only the upper-layer image.

The photographic laboratories at the European Southern Observatory (ESO) are equipped to combine the technique of 'unsharp masking' with that of photographic amplification. This involves making a reversed-film copy of the original plate (the mask). The mask is unsharp, the extent of unsharpness being determined by the thickness of the glass plate. The mask is mounted on the original plate and a high-contrast amplified copy of the plate-mask sandwich is produced in a vacuum contact printer fitted with a diffuse light source. I applied this procedure to deep red IIIaF+RG630 plates of the Rosette nebula NGC2237–2246 that I obtained in January 1990 using the ESO 1-m Schmidt telescope at La Silla, Chile. The F spectroscopic emulsion and RG630 filter combination is useful for studying regions of stellar births, as the peak wavelength, 6,650 Å, is to the red of the H α line.

The enhanced image of plate 8412A is shown in Fig. 1a. At the time of exposure, seeing conditions were at the 1-arcsec level. The enhancement reveals a family of low-surface-brightness cometary globules that cannot be positively identified on R-glass copies of the National Geographic/Palomar Observatory Sky Survey. Because of their semi-parallel tails, the family somewhat resembles the family CG30/31/38 of cometary globules in the Gum nebula⁵, but they are four times farther away, at 1,600 pc (the distance to the Rosette nebula).

Image enhancement alone does not reveal the entire family, because of the very large density range near the optical rim of the nebula. The results of photographic masking followed by image enhancement of plate 8412A are shown in Fig. 1b. At 1,600 pc, 1 arcmin corresponds to 0.47 pc. Figure 1b reveals a further large globule with an apparent foreground star and structural detail in the Rosette nebula (which is totally black in Fig. 1a) as well as the fainter globules. A bright rim around the head is evident. The head is 1.4 arcmin across and contains an estimated 400 $M_{\odot}$ (see below). Also visible is a chain of 'elephant trunk' structures in the southeast quadrant of the Rosette nebula; these have not been reported before (G. Herbig, personal communication); details will be presented elsewhere.

Schneps, Ho and Barrett¹⁰ have studied the 'elephant trunk' globules in the northwest quadrant of the Rosette nebula^{11,12} and conclude that they are most likely the result of stellar wind action in the H II region. Stellar winds sweep up a thin dense shell in the interstellar medium which expands outwards at high

velocities, forming an interstellar bubble. Globules in the shell can be formed either as a result of expansion into an inhomogeneous ambient medium (Pikel'ner¹³), or of Rayleigh-Taylor instabilities in the expansion of the shell, such as that produced by the increase in stellar wind associated with early-type stars.

FIG. 1 *a*, Image amplification of a deep-red Schmidt plate 8412A obtained on 29 January 1990 reveals a new family of cometary globules (circled) with long tails, in the southeast quadrant of the Rosette nebula. The extremely low surface brightness of the globules is evident from the density range of the globules to the Rosette (black). *b*, Image after unsharp masking and image enhancement. Another cometary globule (in the oval) is observed, as well as the family of low-surface-brightness globules seen in *a*. Structural features in the Rosette itself are now also evident. Produced in the ESO darkrooms using a vacuum contact printer and a diffuse light source.

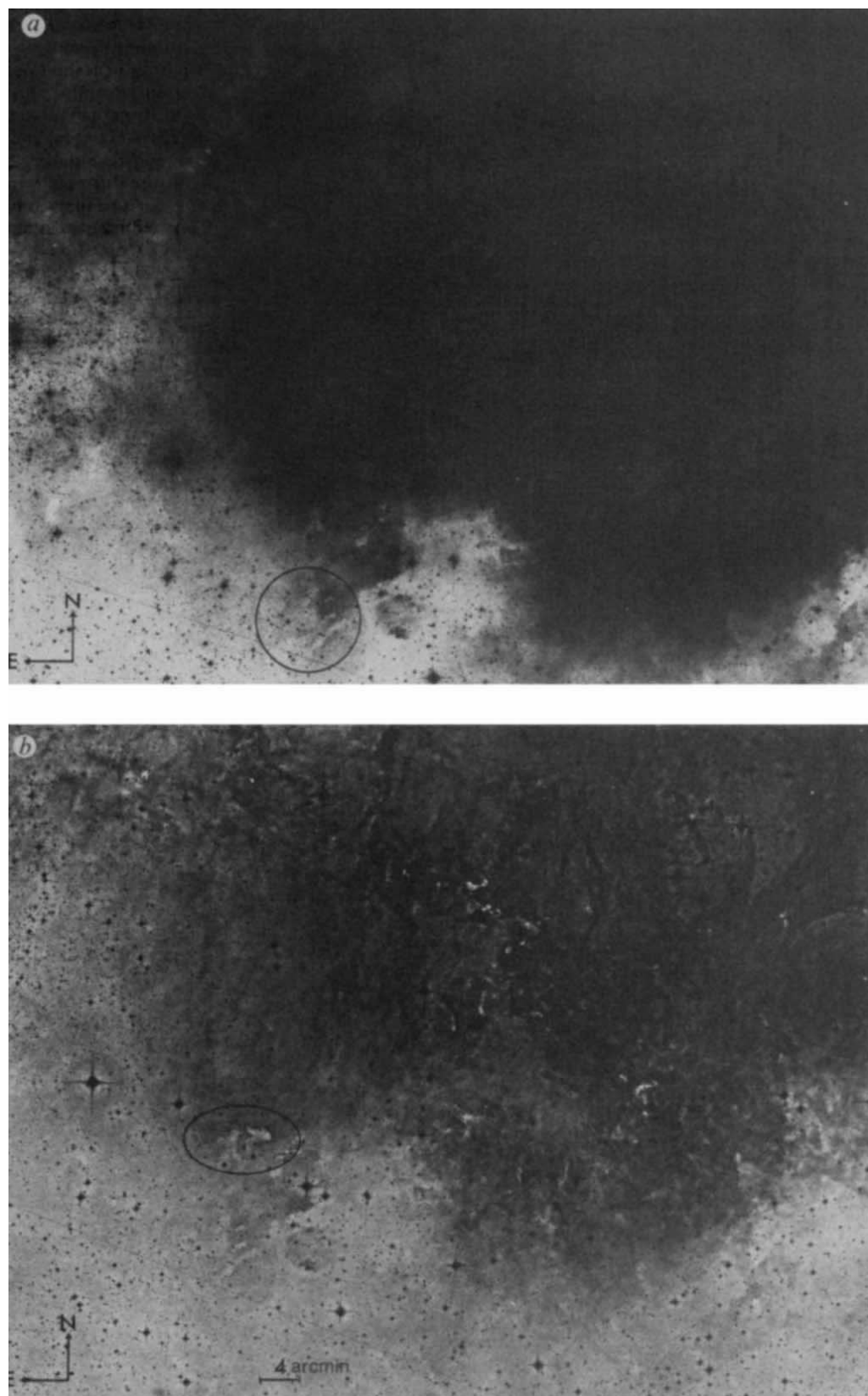
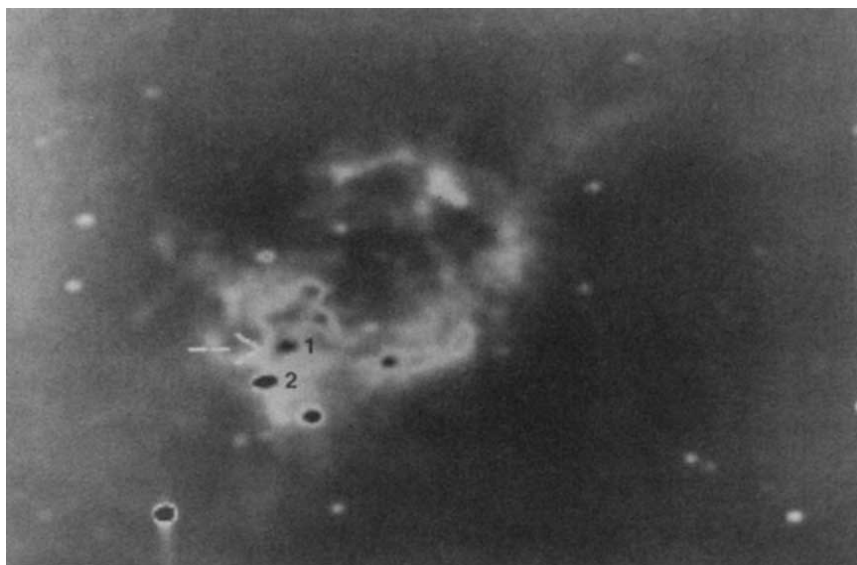


FIG. 2 An IRAS 12- μ m coadded photograph of the Rosette complex. North is up, east to the left. The object '1', marked by an arrow, is the bright source 06314+0427 to which the newly discovered family of cometary globules radially point. Object '2' is AFGL961.



In the latter case, the shell could simply become unstable and break into globules. As stellar wind impinges on the globules, material is driven radially away from the dense heads, forming a luminous tail (see Fig. 6 in ref. 10). The globules slowly evaporate and in such hostile environments most globules probably have lifetimes of a few million years.

The Rosette nebula is the site of spectacular OB star formation; stellar wind from the O stars in the cluster NGC2244, which forms the core of the Mon OB2 association, is believed to be responsible for the expansion of the central cavity¹⁴⁻¹⁷. The model of Schneps *et al.*¹⁰ is consistent with the Rosette nebula being an H II region driven by stellar wind,

expanding at $\sim 20 \text{ km s}^{-1}$.

Schneps *et al.*¹⁰ adopted a particle density of the order of 10^4 cm^{-3} or equivalently $2 \times 10^{-20} \text{ g cm}^{-3}$ for the Rosette globules¹⁰. Assuming that the large globule in Fig. 1b was initially spherical with a radius of 0.7 pc, it would therefore have a mass of $\sim 400 M_{\odot}$. For the family of globules, it would not be unrealistic to adopt a total globular mass of $\sim 500 M_{\odot}$, with most of the mass content being located in the large northern globule.

At the Infrared Processing and Analysis Center (Caltech), I prepared an IRAS photograph of the Rosette complex. The photograph (Fig. 2) shows that the Rosette nebula contains a



FIG. 3 The pointing direction of the globules is indicated by arrows. All globules point radially to the source 06314+0427, coincident with the peak of CO emission and located in the rectangle. The peak¹ lies 15 arcmin northwest of the bipolar infrared source AFGL961 and the reflection nebula²⁶ labelled IRS.

conspicuous infrared source in the southeast, labelled in the IRAS Point Source Catalogue as 06314+0427. The quoted position of this source is right ascension, RA (1950) $6^{\text{h}} 31^{\text{m}} 27^{\text{s}}$, declination, dec (1950), $04^{\circ} 27' 17''$, coincident with the peak of CO emission mapped by Blitz and Thaddeus¹. They found a large (~ 100 -pc major diameter) clumpy molecular-cloud complex associated with the Rosette nebula, with a peak at RA (1950) $6^{\text{h}} 31^{\text{m}} 28^{\text{s}}$, dec (1950) $04^{\circ} 28' 00''$ (uncertainties in RA and dec, ~ 1 arcmin). Peaks of CO emission are excellent tracers of sites of star formation. Figure 3 shows that the globules point radially toward 06314+0427 and the CO peak, which are inside the rectangle. 06314+0427 is located near the Monoceros Ridge, where the ionization front from the Rosette H II nebula is in direct contact with the molecular gas (Fig. 8 in ref 1). It is in this hostile environment that the family of cometary globules have formed.

I used the IRAS pointing history files to secure all the scans by the satellite within 2 arcmin of the target position 06314+0427. The ADDSCAN and SCANPI data-analysis packages were used to secure accurate coadded flux densities of the source at 12, 25, 60 and 100 μm . The resulting flux densities are 20.64, 21.51, 182.01 and 390.34 Jy respectively. The inferred grain temperature at 60–100 μm (deduced from Table 3 in ref. 18) is ~ 36 K, assuming a dust emissivity index of one, indicating a warm infrared source (W. Rice, personal communication).

As the observed flux densities are inversely proportional to the IRAS bandpass frequencies, it is appropriate to use the formulation of Helou *et al.*¹⁸ to calculate the far-infrared flux $\text{FIR} = 1.26(2.58 \text{ FLUXD}(60) + 1.00 \text{ FLUXD}(100)) \times 10^{-14} \text{ W m}^{-2}$ where FLUXD(60) and FLUXD(100) denote the flux densities at 60 and 100 μm , respectively. The FIR parameter can be used for both extragalactic as well as galactic sources with suitable infrared spectra (see appendix in ref. 18), and has been used by Greene and Young¹⁹, for example, in their treatment of dust heating and energy balance in the Rho Ophiuchi Cloud.

The far-infrared luminosity is given by²⁰ $\log L(\text{FIR}) = \log(\text{FIR}) + 2 \log D - 25.48$ where D is the distance (m) and $L(\text{FIR})$ is in units of $L_{\odot}$. For 06314+0427, the observed flux densities yield a lower limit of the total luminosity, $\sim 880 L_{\odot}$. If the material radiating in the far-infrared is only situated on the side of the star toward Earth, the far-infrared energy would greatly underestimate the true radiant luminosity as the remainder would be streaming out unimpeded to give rise to the tails. If the solid-angle covering factor were 1%, the total stellar luminosity would be increased by ~ 100 . If the star is on the zero-age main sequence, the inferred spectral type would then be O7 (ref. 21); using the theory of Weaver *et al.*²², such a star would have more than sufficient energy to expand a $500\text{-}M_{\odot}$ shell to 5–8 pc with an expansion velocity of $\sim 26 \text{ km s}^{-1}$ (a similar expansion velocity was found for globules in the north-west quadrant¹⁰).

The lengths of the tails of the globules circled in Fig. 1a are typically ~ 4 arcmin (or ~ 2 pc), and the globule in Fig. 1b also has a head-tail length of ~ 4 arcmin. Assuming a continuous stellar wind from 06314+0427, the globules in the southeast quadrant are of similar age. (Tail lengths have been used to derive age estimates of globules in, for example, the Gum nebula²³. The inherent assumption is that the flow should not be discontinuous). These globules are much larger compared with the numerous small 'teardrop' globules reported¹² in the northwest quadrant; many of those tiny fragments measure ~ 3 arcsec (~ 0.01 pc) across.

The formation of massive stars has proceeded sequentially in Mon OB2. In accordance with general conclusions drawn by Blaauw²⁴ for many OB associations within 1 kpc of the Sun, the Mon OB2 association contains at least three subassociations whose diameters increase with age¹. The diameter range is ~ 280 pc, for the oldest, to ~ 25 pc, for the young cluster NGC2244. An upper bound of 3×10^6 years for the age of the

third subassociation (NGC2244) can be set by the presence of an O4 star on the main sequence²⁵. According to Blaauw, the youngest of the subassociations is the one most closely associated with the molecular material. From the information for the pre-main-sequence object AFGL961 (Fig. 3), the family of globules pointing to the source 06314+0427 (and which is coincident with the CO peak) could indicate the youngest (fourth) subassociation in Mon OB2, located southeast of the Monoceros Ridge. $\square$

Received 5 June; accepted 16 August 1990.

1. Blitz, L. & Thaddeus, P. *Astrophys. J.* **241**, 676–696 (1980).
2. Malin, D. F. *Nature* **276**, 591–593 (1978).
3. Malin, D. F. *Am. Soc. Photogr. Bull.* **16**, 10–13 (1977).
4. Madsen, C. in *Exploring the Southern Sky* (eds Lausén, S., Madsen, C. & West, R. M.) (Springer-Verlag, New York, 1987).
5. Reipurth, B. *Astr. Astrophys.* **117**, 183–198 (1983).
6. Turner, D. G. *Astrophys. J.* **210**, 65–75 (1976).
7. Heiser, A. M. *Astr. J.* **82**, 973–977 (1977).
8. Perez, M. R., The, P. S. & Westerlund, B. E. *Publ. astr. Soc. Pacif.* **99**, 1050–1066 (1987).
9. Millward, C. G. & Walker, G. A. H. in *The MK Process and Stellar Classification* (ed. Garrison, R. F.) 261–276 (University of Toronto, 1984).
10. Schneps, M. H., Ho, P. T. P. & Barrett, A. H. *Astrophys. J.* **240**, 84–98 (1980).
11. Minkowski, R. *Publ. astr. Soc. Pacif.* **61**, 151–153 (1949).
12. Herbig, G. H. *Publ. astr. Soc. Pacif.* **86**, 604–608 (1974).
13. Pikelner, S. B. *Comments Astrophys.* **6**, 151 (1973).
14. Mathews, W. G. *Astrophys. J.* **144**, 206–215 (1966).
15. Mathews, W. G. *Astrophys. J.* **147**, 965–978 (1967).
16. Smith, M. G. *Astrophys. J.* **182**, 111–120 (1973).
17. Hutchings, J. B. *Astrophys. J.* **203**, 438–447 (1976).
18. Helou, G., Khan, I. R., Malek, L. & Boehmer, L. *Astrophys. J. Suppl.* **68**, 151–172 (1988).
19. Greene, T. P. & Young, E. T. *Astrophys. J.* **339**, 258–267 (1989).
20. Fairclough, J. H. in *Extragalactic Infrared Astronomy, Workshop RAL-85-086* (ed. Gondhalekar, P. M.) (Rutherford Appleton Laboratory, Oxford, 1985).
21. Maeder, A. & Meynet, G. *Astr. Astrophys.* **210**, 155–173 (1989).
22. Weaver, R., McCray, R., Castor, J., Shapiro, P. & Moore, R. *Astrophys. J.* **218**, 377–395 (1977).
23. Hawarden, T. G. & Brand, P. W. J. L. *Mon. Not. R. astr. Soc.* **175**, 19P–22P (1976).
24. Blaauw, A. A. *Rev. Astr. Astrophys.* **2**, 213–246 (1964).
25. Stothers, R. *Astrophys. J.* **175**, 431–452 (1972).
26. Cohen, M. *Astrophys. J.* **185**, L75–L78 (1973).

ACKNOWLEDGEMENTS. I thank Professor H. van der Laan, A. Stockton and H. Ross for invitations to be a Visiting Astronomer at the European Southern Observatory, the Institute for Astronomy (Hawaii) and the Reasons To Believe Institute, and the ESO Director General for allocation of observation time at La Silla. I also thank R. West for permission to work in the ESO Photolabs, C. Madsen for instruction in the darkrooms, and P. Conti, J. Castro, G. Herbig, H-E. Schuster, G. Pizarro, W. Rice, D. Levin and D. Rogstad for their input.

A simple way to make tough ceramics

W. J. Clegg, K. Kendall, N. McN. Alford, T. W. Button & J. D. Birchall

ICI Advanced Materials, PO Box 11, The Heath, Runcorn, Cheshire WA7 4QE, UK

THE major problem with the use of ceramics as structural materials is their brittleness. One way of overcoming this problem is to introduce weak interfaces which deflect a growing crack¹. Polymer composites of this sort can be easily prepared by surrounding fibres with liquid plastic. To make similar structures with ceramic matrices and fibres is difficult and expensive, however, because traditional ceramic processing techniques of powder compaction and sintering prevent densification and cause cracking^{2–4}. Here we describe a simple, inexpensive way of preparing a ceramic material that contains such weak interfaces. Silicon carbide powder is made into thin sheets which are coated with graphite to give weak interfaces and then pressed together and sintered without pressure. Relative to the monolithic material, the apparent fracture toughness for cracks propagating normal to the weak interfaces is increased more than fourfold, and the work required to break the samples increases by substantially more than a hundredfold. The technique should be readily applicable to other ceramics.

By far the most successful method to date for producing high-toughness composites is the infiltration of a ceramic fibre preform with a gaseous ceramic precursor⁵. Several months are

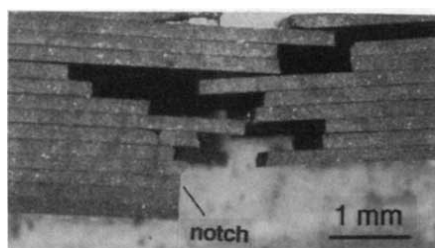


FIG. 1 The fracture surface of a laminate material notched and tested in a three-point bend test. The role of the interface in deflecting cracks can be clearly seen. To the right of the notch the crack was deflected and travelled all the way along the interface to the end of the sample, causing part of the sample to break away.

required to produce parts only a few millimetres thick, however, and several steps of surface grinding are required⁶. The complexity and length of time required for such a process produces materials which, despite their good properties, are very expensive. If ceramic materials are ever to achieve widespread application then some simple and inexpensive fabrication route must be found.

One possibility is to adapt the technology used to make multilayer capacitors. Ceramic powders could be formed into a sheet and the surfaces coated to give suitable interfacial properties. The sheets can then be pressed together to give the desired shape and sintered without pressure. The essential point is that the interface between the sheets is weak enough to deflect cracks and thus increase the resistance to fracture. The process is simple and, as it uses powders as its source of ceramic material, is relatively inexpensive.

To demonstrate the idea, we made thin sheets of silicon carbide powder doped with 0.4 wt% of boron (Superior Graphite, Grade HSC-059s) by mixing the powder with a 40 wt% aqueous solution of polyvinylalcohol acetate (Nippon Gohsei, KH17S) using a high-shear mixer⁷. The boron enables densification of the silicon carbide as described by Prochazka⁸. The resulting dough-like material was pressed into a 2-mm-thick sheet and then rolled into flat sheets ~200 µm thick. The sheets were dried and cut into squares 50 × 50 mm.

The next step was to coat the sheets so as to retain a sufficiently weak interface in the sintered body. We chose graphite as the coating material; it is successfully used as an interfacial layer in existing ceramic composites⁹, possibly because of its low strength in shear, although work by Kendall¹⁰ shows this may not be important in deflecting cracks. Graphite also has the advantage that it does not react with silicon carbide, even at the highest temperatures, as silicon carbon does not form non-

stoichiometric compounds¹¹. The coated sheets were stacked and pressed to form a plaque ~2 mm thick, heated at 1 °C min⁻¹ to 450 °C under argon to pyrolyse the polymer, and then sintered at 2,040 °C under argon for 30 min. The final density was 98% of the theoretical value (3.21 g cm⁻³).

For comparison, monolithic silicon carbide was made by drying some of the 2-mm-thick sheet, pressed from the plastic mix as above, and cutting it into strips ~2 mm wide. The strips were then heated as above to give a body with, again, a final density of 98% of the theoretical.

Both materials had a flexural modulus of 450 GPa. The average three-point-bend strengths were 500 MPa and 633 MPa for the monolithic and laminated material, respectively.

To compare the fracture behaviour of the monolithic and the laminated materials, samples ~2 mm wide were notched perpendicular to the plane of the interfaces and tested in a three-point bend using a span of 40 mm and a cross-head speed of 0.1 mm min⁻¹. As expected the monolithic material behaved in a linear elastic fashion until catastrophic failure occurred at a load σ_c given by the expression¹²

$$\sigma_c = \frac{K_{Ic}}{Y\sqrt{c}} \quad (1)$$

where K_{Ic} is the fracture toughness (the subscript 1 denotes that the mode of loading is tensile and subscript c denotes critical), Y is a geometrical factor for an edge crack in a three-point-bend bar¹³ and c is the notch depth. The fracture toughness calculated from equation (1) was 3.6 MPa m^{-1/2}. The laminated material also deformed in a linear elastic fashion until the crack reached the same stress intensity as the monolith, at which point it began to grow. As soon as the crack reached an interface it was deflected along the interface, preventing catastrophic failure (Fig. 1). In fact, the measured load continued to rise so that the apparent toughness measured using equation (1) reached 15 MPa m^{-1/2}, as shown in Fig. 2. At this point more rapid crack growth occurred and the measured load began to fall, but not catastrophically.

The work required to break the laminated sample was also substantially higher than that in the monolithic material. This work can be obtained from the area under the load-deflection curve, provided that the crack growth is non-catastrophic so that all of the work from the testing machine is used to grow the crack, rather than to provide kinetic energy to the sample^{14,15}. This condition is fulfilled for the laminate material and gives an average work of fracture of 4,625 J m⁻², slightly higher than that of deal wood¹⁵. Values as high as 6,700 J m⁻² were observed in other samples tested. Mecholsky¹⁶ has pointed to difficulties in the measurement of the work of fracture in composite materials once crack deflection or crack branching have occurred. It is precisely because of these effects that composites of

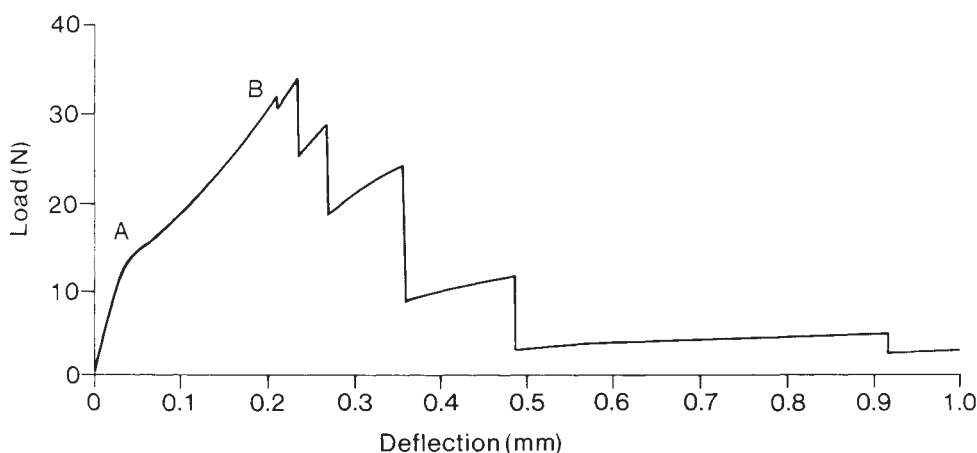


FIG. 2 The load/deflection behaviour of a laminated material notched and tested in a three-point bend test. Crack growth begins where the compliance of the sample first begins to decrease, at A, and is followed by a rising load for further deflection until the point B when crack growth becomes more rapid.

brittle materials are so resistant to crack growth, however, and such measurements give a useful indication of the difficulty of propagating cracks through the material.

The work of fracture of the monolithic material measured from the area under the load-deflection curve is 62 J m^{-2} . In this case, crack growth is unstable and the figure of 62 J m^{-2} is an upper bound to the work required for crack growth. A more accurate estimate for linear elastic materials may be obtained from the fracture toughness¹³ using

$$G = \frac{K_{Ic}^2}{E} (1 - \mu^2) \quad (2)$$

Substituting a value for μ of 0.138 (ref. 17) and taking the measured value of K_{Ic} and E as $3.6 \text{ MPa m}^{-1/2}$ and 450 GPa , respectively, gives a work of fracture of 28 J m^{-2} , close to the range of values $30\text{--}50 \text{ J m}^{-2}$ reported¹⁶. Thus, laminating the material increases the work required to propagate a crack by a factor of over 100.

The fabrication route has potential for other ceramic materials wherever the appropriate powders are available. Other shapes of composite element are also possible. Fibres can be extruded from the plastic mix enabling different composite structures to

be formed. For instance, zirconia powders have been extruded into $150\text{-}\mu\text{m}$ -diameter fibres, dried and pressed together in a small die to form a body with a rope-like structure and then sintered. The resulting material had a work of fracture ten times that of the monolithic material. $\square$

Received 29 June; accepted August 1990.

1. Cook, J. & Gordon, J. E. *Proc. R. Soc. A* **282**, 508–520 (1964).
2. De Jonghe, L. C., Rahman, M. N. & Hsueh, C. H. *Acta Metall.* **34**, 1467–1471 (1986).
3. Clegg, W. J., Alford, N. McN. & Birchall, J. D. *Proc. Brit. Ceram. Soc.* **39**, 247–254 (1987).
4. Clegg, W. J. & Birchall, J. D. *Proc. 4th Int. Conf. on Fibre Reinforced Composites* (ed. Gibson, A. G.) 179 (Mech. Engng Publ., Bury St Edmonds, 1990).
5. Lamicq, P. J., Bernhard, G. A., Dauchier, M. M. & Mace, J. G., *Am. Ceram. Soc. Bull.* **65**, 336–338 (1986).
6. Stinton, D. P., Besmann, T. M. & Lowden, R. A. *Am. Ceram. Soc. Bull.* **67**, 350–355 (1988).
7. Alford, N. McN. Birchall, J. D. & Kendall, K. *Nature* **330**, 51–53 (1987).
8. Prochazka, S. in *Special Ceramics 6* (ed. Popper, P.) 171–182 (Inst. Ceram., Manchester, 1975).
9. Corbin, N. D., Rossetti, G. A. & Hartline, S. D. *Ceram. Engng Sci. Proc.* **6**, 632–645 (1985).
10. Kendall, K. *Proc. R. Soc. A* **344**, 287–302 (1975).
11. Birnie, D. P., Mackrodt, W. C. & Kingery, W. D. *Adv. Ceram.* **23**, 571–584 (1986).
12. Griffith, A. A. *Phil. Trans. R. Soc. A* **221**, 163–198 (1920).
13. Knott, J. F. *Fundamentals of Fracture Mechanics* 134 (Butterworths, London, 1981).
14. Nakayama, J. *Japan J. appl. Phys.* **3**, 422–423 (1964).
15. Tattersall, H. G. & Tappin, G. J. *Mater. Sci.* **1**, 296–301 (1966).
16. Mecholsky, J. J. *Am. Ceram. Soc. Bull.* **65**, 315–322 (1986).
17. *Gmelin Handbook of Inorganic Chemistry, Silicon Supplement* Vol. B2, 95–98 (Springer, Berlin, 1984).

ACKNOWLEDGEMENTS. We thank L. R. Seddon for help with the experiments.

Alkylmercury species in the equatorial Pacific

R. P. Mason & W. F. Fitzgerald

Department of Marine Sciences, University of Connecticut, Avery Point, Groton, Connecticut 06340, USA

HIGH levels of mercury in piscivorous fish constitute a long-standing health hazard^{1–6}. Monomethyl mercury, the main form of mercury in fish, is more toxic than inorganic mercury. But although something is known of the ability of organisms to methylate mercury^{7,8}, the sources, synthesis and fate of methyl mercury in aquatic waters are not well understood. Inorganic and alkylated mercury has been studied in natural waters^{9–11}, precipitation and the atmosphere^{12,13}. We now report evidence of monomethyl and dimethyl mercury in the low-oxygen waters of the equatorial Pacific. The presence of these species has important implications for our understanding of the biogeochemical cycling of mercury in the marine environment. Although the source of monomethyl mercury in open-ocean fish is still unknown, our data show that a pathway exists for the accumulation of methylated mercury in marine pelagic fish.

Recent evidence of significant production and sea-air exchange of mercury in the biologically productive equatorial Pacific Ocean¹⁴, and an indication of a volatile organic mercury species in the sub-thermocline waters¹⁵, prompted this study.

Nine hydrographic stations were occupied in the equatorial Pacific Ocean during the NOAA MB-90-01-RITS cruise of the *Malcolm Baldrige* from 3 January to 18 February 1990. As illustrated in Fig. 1, the cruise consisted of two legs between the Panama Canal and American Samoa. Using protocols¹⁶ designed to avoid trace-metal contamination, water samples were collected in the mixed layer and below the thermocline to a maximum depth of 900 m. Samples were analysed on board for mercury species and for ancillary oceanographic parameters. Air samples were collected continuously, and precipitation was sampled whenever possible.

Dimethyl mercury (DMHg) and total dissolved gaseous mercury (DGM), which consists principally of Hg^0 , and DMHg (ref. 14), were determined within 8 h on separate aliquots taken directly from Teflon-lined Go-Flo bottles. Another aliquot of acidified sea water was analysed for reactive mercury (Hg_R) and monomethyl mercury (MMHg) within 48 h. Reactive or 'easily reducible' mercury is defined as the fraction that is reduced by SnCl_2 at low pH. This fraction includes inorganic complexes, labile organic associations, elemental mercury (Hg^0) and labile particulate mercury¹². Total particulate mercury was also measured.

In general, DMHg and MMHg were found in the waters below the thermocline ($>200 \text{ m}$). Mercury profiles and oceanographic parameters from three representative stations which illustrate the trends and concentration range for all stations are shown in Fig. 2. For the sub-thermocline waters ($>200 \text{ m}$) at all stations, DMHg ranged from $30\text{--}670 \text{ fM}$ (number of determi-

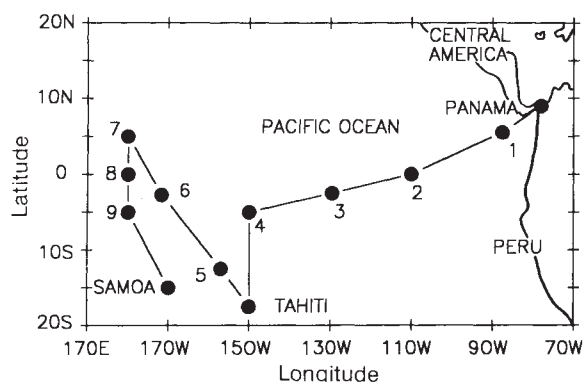


FIG. 1 Path of the *Malcolm Baldrige*. The stations are numbered in sequence of sample collection, as they appear in the text. The cruise consisted of a transect between the Panama Canal and Tahiti (Leg 1), and a track from Tahiti to 5°N , 180°W , then to 5°S , 180°W (Leg 2).

nations, $n = 28$), MMHg from the detection limit to 280 fM ($n = 24$) and Hg_R from 0.72 to 2.85 pM ($n = 30$). Reactive mercury concentrations of 0.47–1.85 pM ($n = 12$) were found in the surface waters (<100 m). The highest levels of Hg_R were in the thermocline region (100–200 m) {1.05–6.9 pM ($n = 12$)}.

At station 4, DGM in the sub-thermocline waters ranged from 280 to 520 fM and DMHg from 160 to 285 fM (Fig. 2). In an earlier study at 150° W, between 10° N and 10° S, Kim and Fitzgerald¹⁵ found concentrations of DGM between 105 and 185 fM at similar depths, and indications of a volatile organic mercury species. The DGM content of surface water (40–325 fM) is also similar to earlier observations (60–225 fM)¹⁴.

The vertical and geographical distributions of the various species provide new insight into the processes controlling the biogeochemical cycling of mercury in the upper ocean. The concentrations of DMHg and MMHg were generally highest around 300–500 m and decreased in the deeper waters and in the thermocline. No detectable DMHg (<5 fM) was found in the surface layer. For MMHg, concentrations at or below the detection limit (50 fM) were generally found in the top 150 m. Stations 2 and 4 were atypical in that detectable MMHg was found in the thermocline zone (Fig. 2). The reactive mercury concentrations were generally highest in the region between 100 and 200 m, which coincided with the thermocline region at most stations. Below 200 m, concentrations of Hg_R were significantly lower than the maximum concentrations. The high sub-thermocline concentrations of DMHg and MMHg fall in

the low-oxygen region, typically between 200 and 500 m, and are associated with low Hg_R . These patterns suggest that DMHg and MMHg are produced in the oxygen-minimum region. We further speculate that the principal substrate for the production of the alkylmercury species is 'labile inorganic mercury'. Analytically, this fraction would be equivalent to the reactive mercury determination adjusted for Hg^0 . Others¹⁷ have also suggested that methylation rate is dependent on the concentration of 'inorganic divalent mercury'. The decrease in Hg_R in the sub-thermocline region that accompanies increasing concentrations of methylated mercury species supports the idea that the methylated species are produced from labile inorganic mercury.

The profiles for both DMHg and MMHg suggest their removal or destruction in the surface waters and production in the sub-thermocline waters. For DMHg, gas exchange is the most likely mechanism, if the flux at the surface is greater than the rate of diffusion across the thermocline. Particulate scavenging in the surface waters may be the primary step leading to removal of MMHg. Circumstantial evidence for this process is indicated by an independent estimate that the MMHg concentration dissolved in surface water is 50 fM (our detection limit), which is the concentration expected to be in equilibrium with the average particulate mercury concentration of 12 pg L⁻¹ measured during this cruise. This calculation was made using literature values for open-ocean particulate concentrations (10 µg L⁻¹) (ref. 18), a MMHg/total Hg ratio for particulate of 0.0088 (ref. 20) and a distribution coefficient for MMHg of 10⁶ (N. Bloom and C. Watras, personal communication). Thus, particulate scavenging

FIG. 2 Profiles of MMHg, DMHg and DGM (left), reactive mercury (middle) and oxygen, nitrate, salinity and temperature (right) at, from top to bottom, stations 2 (0°, 110° W), 4 (5° S, 150° W) and 9 (5° S, 180° W). DMHg was purged from sea water, trapped on a carbon adsorbent, separated using cryogenic gas chromatography, and measured by cold-vapour atomic fluorescence (AFS)³⁰. MMHg was determined in a similar manner following derivitization using sodium tetraethylborate. A pre-extraction step with methylene chloride separates MMHg from chloride ions which interfere with the ethylation⁹. DGM was determined by trapping volatile species on a gold column. Hg_R was separated by $SnCl_2$ reduction, aeration and gold-column trapping¹⁰. The total particulate mercury, filtered from 20-l samples of sea water, was determined by $BrCl$ oxidation of the quartz filters, hydroxylamine neutralization and $SnCl_2$ reduction. AFS was used for all analyses. Procedural blanks were consistent throughout the cruise (DMHg < 5 fM; MMHg, 40 ± 17 fM; DGM, 145 ± 45 fM; Hg_R , 0.58 ± 0.16 pM; particulate Hg, ~100 fM).

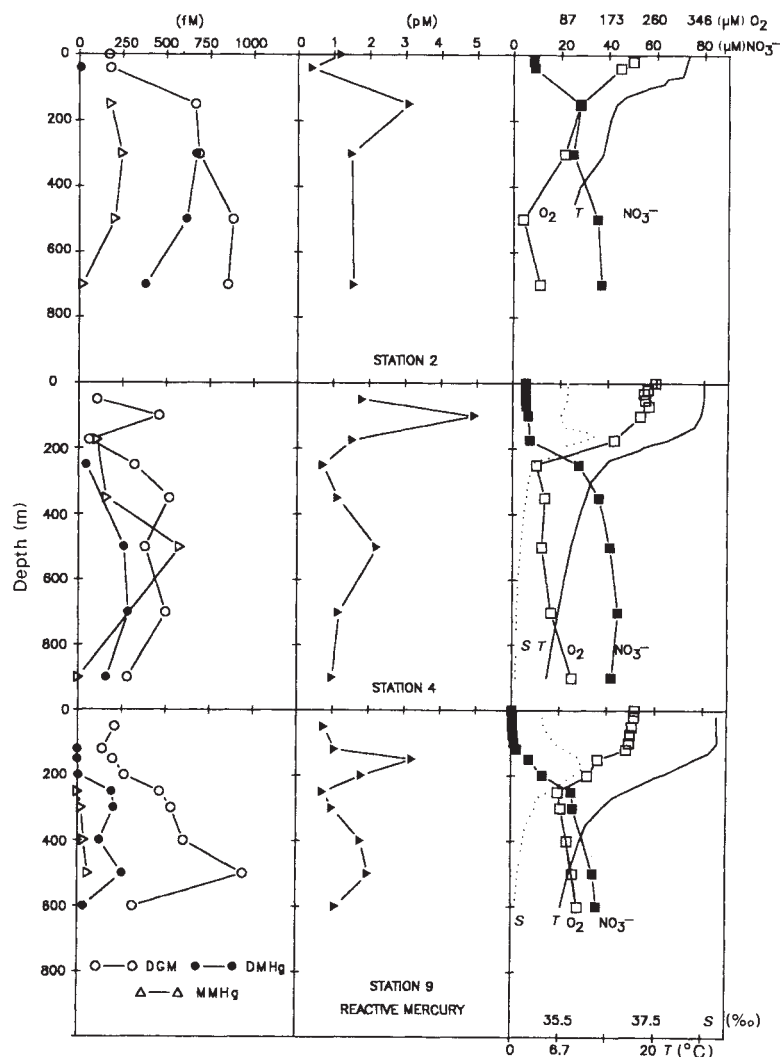
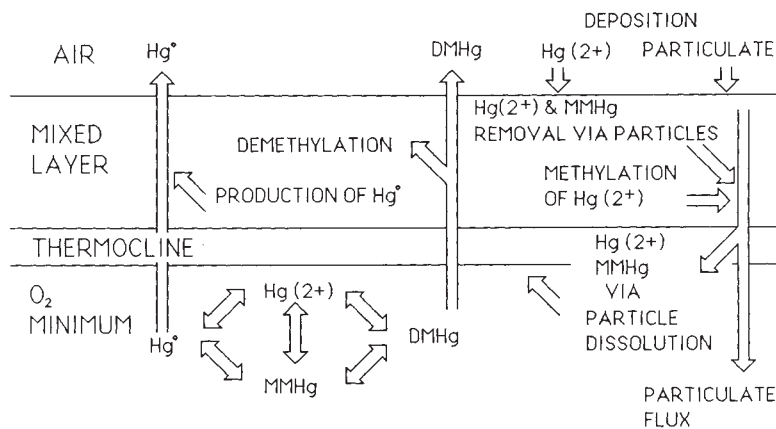


FIG. 3 The preliminary conceptual model displaying the principal cycles of interactions and conversions of the various mercury species in the equatorial Pacific. DMHg and MMHg may be produced in the oxygen-minimum region from labile inorganic mercury substrate, which is supplied to the surface waters by wet and dry deposition, and to the sub-thermocline waters by mixing and by particulate dissolution. Upwelling could supply further Hg(II) to the low-oxygen zone. Horizontal advection, upwelling and diffusion would modify the distributions. Horizontal advection may be important in determining concentrations in the more westerly sub-thermocline waters. Dimethylmercury, lost to the atmosphere by gas exchange through the sea surface, will be rapidly decomposed in the atmosphere³¹. Monomethylmercury is mixed across the thermocline into the surface waters where it is scavenged by particulate adsorption processes. Particulate dissolution in the deeper waters could release bound MMHg and Hg(II) into solution.



could remove MMHg from the surface waters, resulting in undetectable concentrations in the mixed layer.

An east-west decrease in alkylmercury concentration coupled with increasing oxygen concentration was observed. A correlation between DMHg concentration and dissolved oxygen was evident (correlation coefficient $r = 0.55$, probability $p = 0.0004$, $n = 37$); similarly, for MMHg ($r = 0.43$, $p = 0.012$, $n = 35$). A generalized east-west decrease in DGM is also apparent here and in previous work¹⁴.

There was no evidence of alkylmercury species in the atmospheric gaseous fraction. Rain collections were analysed for reactive mercury and MMHg using procedures described above. There was no evidence of MMHg in rain (< 50 fM), contrary to what we have found for continental precipitation^{13,19}. Thus, the atmosphere is not a significant source of MMHg to the open ocean. This observation is corroborated by our studies and those of our colleagues (C. Watras, N. Bloom and J. Hurley, personal communication) in the lakes of northcentral Wisconsin where we find in-lake processes to be responsible for a large proportion of the MMHg found in fish¹⁹.

Although both biotic and abiotic mechanisms for alkylation of mercury in both oxic and anoxic environments have been postulated^{7,8}, biological methylation in low-oxygen environments is likely to be the most significant process²¹. Little information on the stability and formation mechanisms of DMHg is available. A consideration of the chemical thermodynamics²² for various reactions of alkylmercury species indicates that interaction of Hg(II) , or MMHg, with a methyl carbanion is highly favoured, yielding MMHg and DMHg, respectively. Such reactions, for example, are likely between mercury species and methylcobalamin²³. DMHg is a product of the reaction between

Hg(II) and methylcobalamin, especially under mildly reducing conditions²⁴. Under conditions of excess Hg(II) , MMHg is the main product²³. DMHg formation in the presence of dead fish tissue has been demonstrated²⁵, and DMHg is produced in alkaline anoxic sediments¹⁷. Compeau and Bartha²⁶ showed that DMHg is produced in anoxic estuarine sediments spiked with MMHg, and that demethylation of MMHg to Hg^0 occurs in both oxic and anoxic waters. Production of MMHg has been demonstrated under a wide variety of conditions, and by numerous organisms⁸. A recent report²⁷ suggests that sulphate-reducing bacteria are the principal methylators of mercury in anoxic estuarine sediments. In addition, methylation in the water columns of coastal marine surface waters²⁰ and fresh water^{28,29} has been found. In ocean waters, methylation of Hg(II) and Hg^0 by methyl donor compounds (such as methyl iodide)²³ is also a potential source of alkylmercury species. Demethylation could account for the loss of DMHg in the more oxygenated thermocline waters.

Although more research is needed to investigate the importance of the various formation and decomposition mechanisms, our results indicate that there is *in situ* production of alkylmercury species in the open ocean and that mercury may accumulate through the food chain into pelagic marine fish. We propose a preliminary conceptual model, outlined in Fig. 3, for the cycling of mercury species in open-ocean equatorial waters. We suggest that labile inorganic mercury is the most probable substrate for methylation and that a low-oxygen environment is conducive to the formation of alkylmercury species. This model provides a working framework for further investigations of alkylated mercury compounds in the marine environment. □

Received 11 June; accepted 15 August 1990.

- Huckabee, J. W., Elwood, J. W. & Hilderbrand, S. G. in *The Biogeochemistry of Mercury in the Environment* (ed. Nriagu, J. O.) 277–302 (Elsevier, Amsterdam, 1979).
- Windom, H. L. & Kendall, D. R. in *The Biogeochemistry of Mercury in the Environment* (ed. Nriagu, J. O.) 312–323 (Elsevier, Amsterdam, 1979).
- Wiener, J. G. & Stokes, P. M. *Environ. Toxicol. Chem.* **9**, 821–823 (1990).
- Tsubaki, T. & Irukayama, K. *Minamata Disease* (Elsevier, Amsterdam, 1977).
- Takizawa, Y. in *The Biogeochemistry of Mercury in the Environment* (ed. Nriagu, J. O.) 325–367 (Elsevier, Amsterdam, 1979).
- Bakir, F. et al. *Science* **181**, 230–241 (1973).
- Beijer, K. & Jernelov, A. in *The Biogeochemistry of Mercury in the Environment* (ed. Nriagu, J. O.) 203–210 (Elsevier, Amsterdam, 1979).
- Robinson, J. B. & Tuovinen, O. H. *Microbiol. Rev.* **48**, 95–124 (1984).
- Bloom, N. *Can. J. Fish. Aquat. Sci.* **46**, 1131–1140 (1989).
- Gill, G. A. & Fitzgerald, W. F. *Mar. Chem.* **20**, 227–243 (1987).
- Lee, Y. H. & Mowrer, J. *Anal. chim. Acta* **221**, 259–268 (1989).
- Fitzgerald, W. F. in *Chemical Oceanography* (eds Riley, J. P. & Chester, P.) Vol. 10, 152–187 (Academic, London, 1989).
- Bloom, N. & Watras, C. *Sci. tot. Envir.* **87/88**, 199–207 (1990).
- Kim, J. P. & Fitzgerald, W. F. *Science* **231**, 1131–1133 (1986).
- Kim, J. P. & Fitzgerald, W. F. *Geophys. Res. Lett.* **15**, 40–43 (1988).
- Gill, G. A. & Fitzgerald, W. F. *Deep Sea Res.* **32**, 287–297 (1985).

- Fagerstrom, T. & Jernelov, A. *Wat. Res.* **6**, 1193–1202 (1972).
- Sackett, W. H. in *Chem. Oceanography* (eds Riley, J. P. & Chester, P.) Vol. 7, 127–172 (Academic, London 1978).
- Fitzgerald, W. F., Vandal, G. M. & Mason, R. P. (abstr.) *Proc. 10th Soc. envir. Tox. Chem. Mtg* (1989).
- Topping, G. & Davies, I. M. *Nature* **290**, 243–244 (1981).
- Olson, B. H. & Cooper, R. C. *Wat. Res.* **10**, 113–116 (1976).
- Stumm, W. & Morgan, J. J. *Aquatic Chemistry* (Wiley, New York, 1981).
- Craig, P. J. in *Organometallic Compounds in the Environment* (ed. Craig, P. J.) 65–110 (Wiley, New York, 1986).
- Wood, J. M., Kennedy, F. S. & Rosen, C. G. *Nature* **220**, 173–174 (1968).
- Jensen, S. & Jernelov, A. *Nature* **223**, 753–754 (1969).
- Compeau, G. C. & Bartha, R. *Appl. envir. Microbiol.* **48**, 1203–1207 (1984).
- Compeau, G. C. & Bartha, R. *Appl. envir. Microbiol.* **50**, 498–502 (1985).
- Xun, L., Campbell, N. E. R. & Rudd, J. W. M. *Can. J. Fish. Aquat. Sci.* **44**, 750–757 (1987).
- Furutani, A. & Rudd, J. W. M. *Appl. envir. Microbiol.* **40**, 770–776 (1980).
- Bloom, N. & Fitzgerald, W. F. *Anal. chim. Acta* **208**, 151–161 (1988).
- Niki, H., Maker, P. D., Savage, C. M. & Breitenbach, L. P. *J. phys. Chem.* **24**, 3722–3724 (1983); **24**, 4978–4981 (1983).

ACKNOWLEDGEMENTS. We thank the captain, officers, crew and survey technicians of the *Malcolm Baldrige* for their assistance, and G. Harvey for the opportunity to participate in the cruise. This research was funded by the Research Foundation of the University of Connecticut.

The influence of solar forcing trends on global mean temperature since 1861

P. M. Kelly & T. M. L. Wigley

Climatic Research Unit, School of Environmental Sciences,
University of East Anglia, Norwich NR4 7TJ, UK

It has been suggested^{1–5} that solar variability may be a significant contributor to the long-term warming trend that has been observed in global mean surface air temperature since the mid-nineteenth century. Here we consider the role of long-term solar irradiance changes associated with sunspot fluctuations, extending an earlier analysis by Reid⁴. We find that solar variability is unlikely to have accounted for more than a small fraction of the observed warming.

Reid used an upwelling–diffusion energy-balance climate model^{6,7} to determine the best fit between the record of sunspot numbers, used as a proxy for variations in solar irradiance, and the record of global mean sea surface temperature derived by Folland *et al.*⁸, both with and without forcing owing to enhancement of the greenhouse effect. A simple relationship between sunspot number and long-term variations in the solar constant S was assumed

$$S_t = S_0[1 + \alpha(N_t - N_0)]$$

where N_t is the sunspot number at time t after application of an 11-year running mean and S_0 and N_0 are reference values. The constant α was varied to identify a subjective best fit between the amplitudes of the modelled and observed temperature changes. Reid presented results for only one value of the climate sensitivity, corresponding to an increase of 2.5 °C in the equilibrium global mean temperature for a doubling of the atmospheric concentration of carbon dioxide. We have extended his analysis by considering a range of climate sensitivities and by identifying the best fit in a quantitative fashion.

To assess how global mean temperature should respond to variations in forcing, we used a model of the climate system⁹ that is conceptually the same as the model used by Reid. The model takes into account the exchange of heat between land and sea, atmosphere and ocean and between the hemispheres, and incorporates an upwelling–diffusive parametrization of ocean mixing which enables it to simulate the transient response of the climate system to time-dependent changes in forcing. The model parameters used were: ocean diffusivity, 1.0 cm² s^{–1}; upwelling velocity, 4 m yr^{–1}; depth of the ocean mixed layer, 100 m; temperature of sinking water, constant. Sensitivity studies indicate that our conclusions are not dependent on these assumptions.

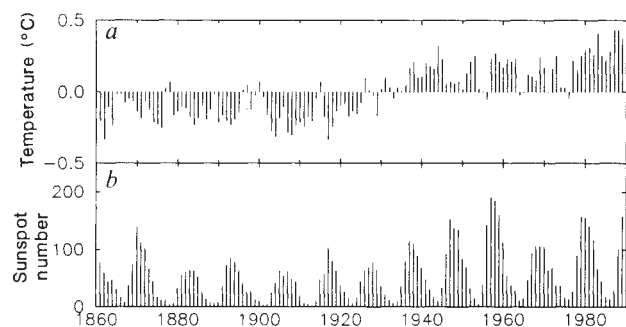


FIG. 1 a, Global mean surface air temperature, annual values expressed as departures from the 1861–1989 mean, and b, annual mean sunspot numbers.

TABLE 1 Best-fit combinations of ΔT_{2x} and β , with error variance

Case	Forcing	Climate sensitivity ΔT_{2x} (°C)	Solar forcing factor, β	Error variance (%)
1	Solar alone	5.0	0.012	44.3
2	Solar alone	1.0	0.038	45.1
3	Greenhouse alone	1.5	0.000	40.2
4	Solar and greenhouse	1.0	0.016	35.0
5	Solar and greenhouse	1.5	0.004	38.4

The error variance is the ratio of the variance remaining after subtracting the modelled from the observed temperatures to the original variance in the observed data.

We carried out two types of experiments. First, it was assumed that solar variability alone is responsible for the observed fluctuations in global mean temperature. Following Reid, changes in solar radiative forcing, ΔQ , were assumed to be related to changes in sunspot number, ΔN , by $\Delta Q = \beta \Delta N$, and the best-fit value of β , the solar-forcing scaling factor, was identified for a range of climate sensitivity values. As there is no doubt that the marked rise in greenhouse-gas concentrations over the past few centuries has had a direct radiative forcing effect, it is, in fact, illogical to consider solar effects in isolation. This point has been noted by Hansen and Lacis¹⁰ and, indeed, by Reid himself, although it has been ignored by some influential commentators (see, for example, ref. 5). We considered this first possibility only for the sake of completeness. Our second experiment combines more realistically the forcing owing to enhancement of the greenhouse effect and a solar contribution that varies with β .

In both experiments, the solar-forcing scaling factor was varied between –0.02 and 0.04 in increments of 0.002. Recent satellite observations suggest that β is about +0.001 (see, for example, refs 11 and 12). This value results in a negligible solar effect on global mean temperature¹³. This estimate of β was derived from observations taken during a limited period of the 11-year sunspot cycle, however, and may not be representative of longer-term relationships. For example, low-frequency variations in the solar radius, as well as leading to irradiance changes (as assumed here), may also play a part in explaining the modulation of the 11-year sunspot cycle. This could result in a nonlinear sunspot–irradiance relationship which would not be apparent in the short satellite record. There is some evidence of a link between global climate change and solar variability on century and longer timescales which would imply a greater value of β at lower frequencies (see, for example, ref. 14). Although rather speculative, this possibility cannot be ruled out.

The climate sensitivity is specified in terms of the change in the equilibrium global mean temperature for a carbon dioxide doubling, ΔT_{2x} . This was varied between 1.0 and 5.0 °C in increments of 0.5 °C, spanning the current range of uncertainty¹⁵. For greenhouse forcing, we used concentration histories and concentration–forcing relationships for the most important greenhouse gases (carbon dioxide, chlorofluorocarbons, methane and nitrous oxide) recommended by Working Group 1 of the Intergovernmental Panel on Climate Change¹⁵. The simulations extend from 1765 to 1989. As an illustration, the greenhouse forcing changes in this period were: 1765–1900, 0.53 W m^{–2}; 1900–60, 0.64 W m^{–2}; 1960–90, 1.28 W m^{–2} (1990 value extrapolated from 1989).

The best fit between modelled and observed temperature was determined, for the values of β and ΔT_{2x} specified earlier, by the root-mean-square error, E_{rms}

$$E_{rms} = \sqrt{\frac{1}{N} \sum (O_i - M_i)^2}$$

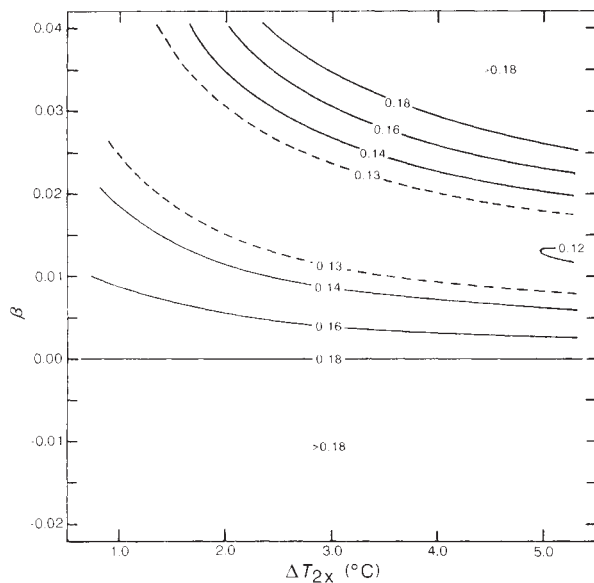


FIG. 2 Variation of the root-mean-square error, E_{rms} , with the solar-forcing scaling factor, β , and the climate sensitivity, ΔT_{2x} , for the case of solar forcing alone. Contours are not plotted when E_{rms} exceeds the standard deviation of the observed data (0.18 °C).

where O_i and M_i are departures in the observed and modelled yearly temperatures from their respective 1861–1989 means and N is the number of annual observations ($N = 139$).

As an indicator of global mean temperature, Reid⁴ used the sea-surface-temperature record, as his earlier work had indicated an apparent correlation between sunspot numbers and tropopause height over the western tropical Pacific¹⁶. There is, however, no reason to expect that there should be any difference between the response of land and ocean temperature to variations in solar irradiance on decadal and longer timescales¹⁷. We used the latest estimate of global mean surface air temperature derived from both land and marine observations¹⁵ (Fig. 1). This record combines an average of the most recent revisions of the marine records of Folland *et al.*⁸ and Jones *et al.*¹⁸ (both of which differ noticeably from the earlier marine record of Folland *et al.* used by Reid) with the up-dated land record of Jones *et al.*^{19,20}. Our results are based on analysis of the complete record, extending from 1861 to 1989.

To test for reproducibility, we repeated the analysis on the more reliable subset of this record, which covers the period 1905–89, and on the record of global mean sea surface temperature used by Reid; there were no significant differences between the results. We have also used both a low-pass filtered sunspot record (following Reid, who used an 11-year running mean to define long-term variations in the record) and the unfiltered record. The following discussion is based on analysis of the filtered data to ensure compatibility with Reid; reference is made to any substantial differences in the results based on the unfiltered data.

Figure 2 shows the variation of the root-mean-square error E_{rms} with β and ΔT_{2x} when solar forcing alone is specified. For $\Delta T_{2x} = 2.5$ °C, E_{rms} reaches a minimum at $\beta = 0.019$. This is in broad agreement with Reid's findings (for his best-fit value of α , 1.08×10^{-4} , $\beta = 0.025$). But somewhat better fits (lower E_{rms}) can be obtained for larger values of ΔT_{2x} , requiring progressively smaller values of β . In the range of ΔT_{2x} values considered, the lowest E_{rms} is reached when $\Delta T_{2x} = 5.0$ °C and $\beta = 0.012$. Although this could be judged to be an optimum solution, it is difficult to justify for two reasons. First, this value of β is more than ten times that suggested by recent satellite observations. Second, such a high climate sensitivity would have to have produced a large greenhouse warming. Suboptimal solutions

for lower ΔT_{2x} are close to this optimal solution, but these are also unlikely because, as ΔT_{2x} decreases, the implied value of β rises, becoming increasingly different from the value indicated by the satellite data. When unfiltered, rather than filtered, sunspot data are used, E_{rms} increases considerably for ΔT_{2x} values towards the lower end of the range.

We now consider the more plausible hypothesis that the observed trend in global mean temperature is explained by a combination of greenhouse and solar forcing. Figure 3 shows the variation of E_{rms} with β and ΔT_{2x} for this case. With zero solar forcing ($\beta = 0.0$), E_{rms} is at a minimum when $\Delta T_{2x} = 1.5$ °C. For most ΔT_{2x} values, the addition of a positive solar contribution ($\beta > 0.0$) results in a higher E_{rms} than is the case when the solar component is zero. The only improvement in fit for a positive solar contribution occurs for $\Delta T_{2x} < 2.0$ °C. E_{rms} is at a minimum when $\beta \approx 0.004$ for $\Delta T_{2x} = 1.5$ °C, and when $\beta = 0.016$ for $\Delta T_{2x} = 1.0$ °C. These values of β are four and sixteen times, respectively, that suggested by the satellite data. This discrepancy must be explained if the solar contribution is to be justified. In any event, the inclusion of a non-zero solar forcing term in these cases results in little improvement in E_{rms} . In terms of percentage variance explained, the improvement amounts to at most 5% of the variance of the observed data (Table 1; compare also cases 3, 4 and 5 in Fig. 4). This is much less than implied by Reid⁴ in his analysis of combined greenhouse and solar forcing.

Our results contradict Reid's assertion that a reasonable fit can be obtained using both greenhouse and solar forcing with $\beta = 0.017$ (equivalent to his optimal value of α for combined forcing) and $\Delta T_{2x} = 2.5$ °C. For this climate sensitivity, our best fit occurs when $\beta = -0.004$, implying that higher sunspot numbers are associated with cooling not warming. In fact, the lowest E_{rms} consistently occurs when β is negative for $\Delta T_{2x} > 2.0$ °C (Fig. 3), suggesting that low-frequency solar variability may have offset greenhouse warming over the past 100 years.

Our findings should not be considered definitive proof that enhancement of the greenhouse effect has accounted for a substantial proportion of the warming trend observed over the past 100 years. Nor should too much weight be attached to the best-fit value of the climate sensitivity ($\Delta T_{2x} = 1.5$ °C) identified here for greenhouse forcing alone. To isolate the greenhouse signal

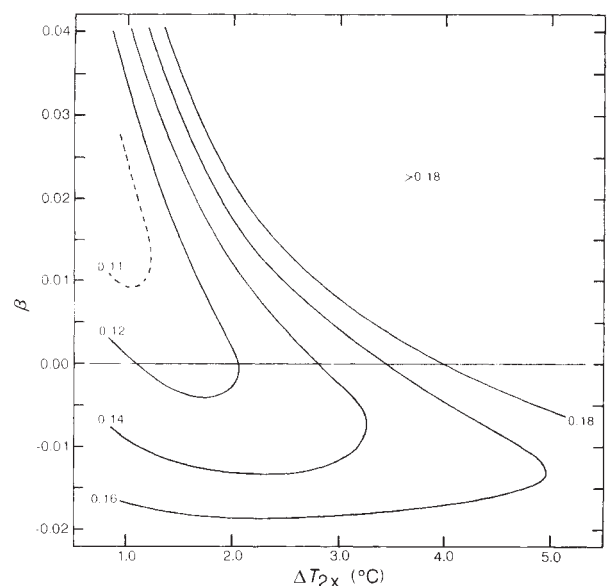


FIG. 3 Variation of the root-mean-square error, E_{rms} , with the solar-forcing scaling factor, β , and the climate sensitivity, ΔT_{2x} , for the case of greenhouse and solar forcing. Contours are not plotted when E_{rms} exceeds the standard deviation of the observed data (0.18 °C).

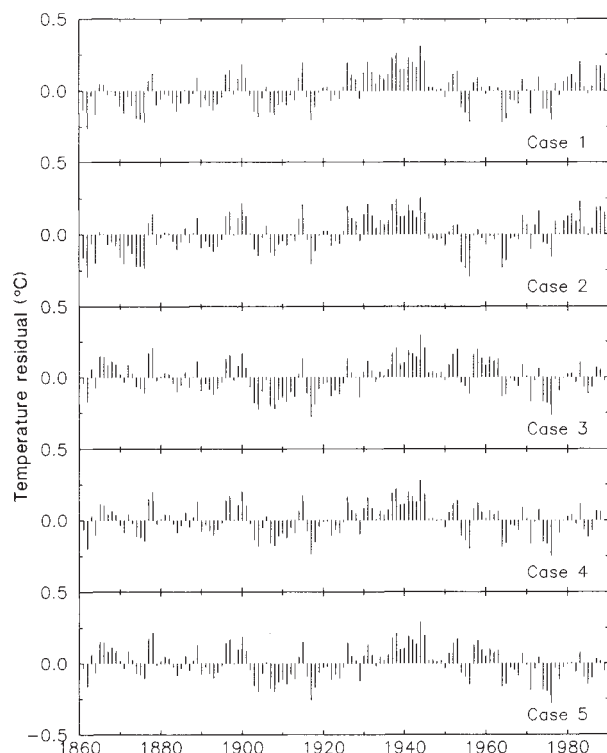


FIG. 4 The difference between the observed and modelled global mean temperature for the five cases cited in Table 1. The cases which include greenhouse forcing, cases 3, 4 and 5, result in lower error variance than the cases based on solar forcing alone, cases 1 and 2. On the basis of error variance, case 4 is the most plausible simulation although β in this case is considerably higher than suggested by recent satellite data. Not surprisingly, fluctuations on interannual to decadal timescales are present in the residuals in each case. None of the simulations adequately accounts for the warmth of the years around 1940 and the trends in adjacent decades. This discrepancy may be the result of the influence of other forcing agents, internal climate variability and/or other factors neglected in these simulations.

and determine the climate sensitivity, other factors, such as internally generated climate variability²¹ and the effect of additional external forcing agents¹⁻³, would have to be taken into account²².

Our results do indicate that, since the mid-nineteenth century, the influence of the enhanced greenhouse effect on global mean temperature has almost certainly dominated over the direct influence of solar variability. For solar variability to have been an important contributing factor, past low-frequency variations in solar irradiance would have to have been far greater than recent satellite studies indicate. There is no firm evidence to support this contention. Compared to the well established observational and theoretical basis for the influence of increased greenhouse forcing, the case for a solar effect on recent changes in global climate is speculative in the extreme. □

12. Foukal, P. & Lean, J. *Science* **247**, 556-558 (1990).
13. Wigley, T. M. L. & Raper, S. C. B. *Geophys. Res. Lett.* (in the press).
14. Wigley, T. M. L. & Kelly, P. M. *Phil. Trans. R. Soc. A* **330**, 547-560 (1990).
15. Intergovernmental Panel on Climate Change *Scientific Assessment of Climate Change, Working Group 1 Report* (Intergovernmental Panel on Climate Change/World Meteorological Organization/United Nations Environment Programme, Geneva, 1990).
16. Gage, K. & Reid, G. C. *Geophys. Res. Lett.* **8**, 187-190 (1981).
17. Wigley, T. M. L., Angell, J. K. & Jones, P. D. in *Detecting the Climatic Effects of Increasing Carbon Dioxide*, DOE/ER-0235 (eds MacCracken, M. C. & Luther, F. M.) 55-90 (US Department of Energy, Washington, DC, 1985).
18. Jones, P. D., Wigley, T. M. L. & Wright, P. B. *Nature* **322**, 430-434 (1986).
19. Jones, P. D. *et al.* *J. Clim. appl. Met.* **25**, 161-179 (1986).
20. Jones, P. D., Raper, S. C. B. & Wigley, T. M. L. *J. Clim. appl. Met.* **25**, 1213-1230 (1986).
21. Wigley, T. M. L. & Raper, S. C. B. *Nature* **344**, 324-327 (1990).
22. Wigley, T. M. L. & Barnett, T. P. in *Scientific Assessment of Climate Change, Working Group 1 Report* (Intergovernmental Panel on Climate Change/World Meteorological Organization/United Nations Environment Programme, Geneva, 1990).

ACKNOWLEDGEMENTS. This work was supported by the US Department of Energy, Carbon Dioxide Research Division.

A carbon isotope record of CO₂ levels during the late Quaternary

John P. Jasper & J. M. Hayes

Biogeochemical Laboratories, Departments of Geological Sciences and of Chemistry, Indiana University, Bloomington, Indiana 47405-5101, USA

ANALYSES of gases trapped in continental ice sheets have shown that the concentration of CO₂ in the Earth's early atmosphere increased from 180 to 280 p.p.m. during the most recent glacial-interglacial transition¹. This change must have been driven by an increase in the concentration of CO₂ dissolved in the mixed layer of the ocean². Biochemical and physiological factors associated with photosynthetic carbon fixation in this layer should lead to a relationship between concentrations of dissolved CO₂ and the carbon isotopic composition of phytoplanktonic organic material³, such that increased atmospheric CO₂ should enhance the difference in ¹³C content between dissolved inorganic carbon and organic products of photosynthesis. Here we show that a signal related to atmospheric CO₂ levels can be seen in the isotope record of a hemipelagic sediment core, which we can correlate with the CO₂ record of the Vostok ice core. Calibration of the relationship between isotope fractionation and CO₂ levels should permit the extrapolation of CO₂ records to times earlier than those for which ice-core records are available.

Popp and coworkers³ have suggested a relationship between carbon isotope fractionation and dissolved CO₂ of the form

$$\epsilon_P = a \log c + b \quad (1)$$

where ϵ_P is the isotope effect (‰) associated with photosynthetic fixation of carbon, c is the concentration of dissolved CO₂ (μM), and a and b are constants. In an open system, the isotope fractionation between substrate and product is numerically equal to ϵ . The functional form of equation (1) seems to be representative of biologically mediated ϵ_P -CO₂ relationships, but the dependence of a and b on physiological, ecological and environmental factors is just beginning to be characterized³⁻⁵. Here, this problem is explored by comparison of historical records of ϵ_P and p_{CO_2} .

To determine values of ϵ_P in ancient environments, we must reconstruct the isotopic compositions of dissolved CO₂ and phytoplanktonic biomass. In surface sea water, dissolved CO₂ is depleted in ¹³C relative to total dissolved inorganic carbon (DIC) by 8.8‰ ($T = 24^\circ\text{C}$, $\text{pH} = 8.2$). In turn, calcite tests produced by the planktonic foraminifera *Globigerinoides ruber* are depleted in ¹³C relative to DIC by an average of 0.5‰^{6,7}. Determination of the isotopic composition of these microfossils ($\delta_{G. ruber}$) therefore allows estimation of the ¹³C content of ancient dissolved CO₂. Although all sedimentary organic carbon is ultimately of photosynthetic origin, most of it has been reworked and isotopically fractionated during its passage through the marine food chain⁸. To construct a record of isotopic

Received 21 May; accepted 17 August 1990

1. Hansen, J. *et al.* *Science* **213**, 957-966 (1981).
2. Gilliland, R. L. *Clim. Change* **4**, 111-131 (1982).
3. Gilliland, R. L. & Schneider, S. H. *Nature* **310**, 38-41 (1984).
4. Reid, G. C. *Nature* **329**, 142-143 (1987).
5. Seitz, F., Jastrow, R. & Nierenberg, W. A. *Scientific Perspectives on the Greenhouse Problem* (George C. Marshall Institute, Washington, DC, 1989).
6. Hoffert, M. I., Callegari, A. J. & Hsieh, C.-T. *J. geophys. Res.* **85**, 6667-6679 (1980).
7. Hoffert, M. I. & Flannery, B. P. in *Projecting the Climatic Effects of Increasing Carbon Dioxide*, DOE/ER-0237 (eds MacCracken, M. C. & Luther, F. M.) 149-190 (US Department of Energy, Washington, DC, 1985).
8. Folland, C. K., Parker, D. E. & Kates, F. E. *Nature* **310**, 670-673 (1984).
9. Wigley, T. M. L. & Raper, S. C. B. *Nature* **330**, 127-131 (1987).
10. Hansen, J. & Lacis, A. *Nature* **346**, 713-719 (1990).
11. Lean, J. & Foukal, P. *Science* **240**, 906-908 (1988).

compositions of microalgae, it is therefore necessary to isolate specific molecular products of these organisms³. We have used isotope-ratio-monitoring gas chromatography/mass spectroscopy techniques⁹ to determine isotopic compositions of C_{37} alkadienones ($\delta_{37:2}$), which can be readily isolated from many sediments¹⁰. These molecules derive specifically from prymnesiophyte algae (such as coccolithophorids)¹¹. Analyses of laboratory cultures show that the alkenones are depleted in ^{13}C by 3.8‰ relative to total algal cell material. Values for ϵ_P are determined from

$$\epsilon_P = 1,000[(\delta_P + 1,000)/(\delta_d + 1,000) - 1] \quad (2)$$

where δ_P (‰) ($=\delta_{37:2} + 3.8$) and δ_d (‰) ($=\delta_{G. ruber} - 8.3$) are the carbon isotopic compositions of prymnesiophyte biomass and dissolved CO_2 , respectively.

We have examined samples from a sediment core obtained at Deep Sea Drilling Project site 619, in the Pigmy Basin, northern Gulf of Mexico^{12,13}. Bottom waters in the basin are oxic and of normal salinity¹³. Sedimentation rates determined for oxygen isotope substages 1 to 5b^{14,15} varied from 71 to 92 cm kyr⁻¹ during interglacial periods, and from 230 to 300 cm kyr⁻¹ during glacial periods, presumably owing to

increased erosion of continental material during times of falling sea level and at low sea-level stands^{14,16,17}. Individual samples spanned between 17 and 70 years and were also mixed by bioturbation, thus averaging effects of seasonal variations in p_{CO_2} .

The $\delta_{37:2}$, $\delta_{G. ruber}$ and ϵ_P records are shown in Fig. 1. The data points marked by open circles are believed to derive from turbidity flows or sediment slumps, common to the northern Gulf of Mexico¹⁸. They have been recognized on the basis of discontinuities in the records of ^{13}C in total organic carbon¹⁵, ^{18}O in carbonates¹⁴, and U_{37}^K ($C_{37:2}/(C_{37:2} + C_{37:3})$, where $C_{37:2}$ and $C_{37:3}$ are the concentrations of the di- and tri-unsaturated alkenones; see ref. 10). In spite of the 2,300-m depth of the basin, the deepest portion of the core (termed isotope substage '5c') contained shallow-water microfossils and quartz sand, indicating the presence of displaced sediments^{14,19}.

The chronology of the Vostok ice core is based on a two-dimensional snow accumulation and ice-flow model²⁰, whereas that of marine sediments is commonly based on the oxygen isotope record in carbonate microfossils. Precise correlation (± 2 kyr) between ice and sediment cores can be difficult. Moreover, whereas p_{CO_2} values have been directly recorded in the ice by physical entrapment of an air sample, the recording of CO_2 levels in sediment cores is indirect. Variable factors of two kinds are involved. First, the concentration of CO_2 dissolved in the photic zone may not be in equilibrium with atmospheric p_{CO_2} (ref. 21). Second, relationships between ϵ_P and c are only now being quantified. We will consider these points separately.

Concentrations of dissolved CO_2 in surface waters of the Gulf of Mexico 2–100 kyr ago are not directly known. If we use the Vostok record of atmospheric p_{CO_2} values¹ to estimate values of c and assume equilibration with the atmosphere at 25 °C during the Holocene and at 23 °C during the glacial interval¹⁰, regression of ϵ_P on $\log c$ for the points at 2.9, 8.5, 26.7, 33.9, 46.2, 60.1, 80.2 and 91.8 kyr BP yields $a = -32.9$, $b = 14.3$ (the calibration data are points for which stratigraphic uncertainties correspond to Vostok p_{CO_2} ranges of ≤ 20 μatm). Using these constants to calculate dissolved CO_2 concentrations corresponding to each ϵ_P value, and converting the concentrations to equilibrium p_{CO_2} values, the isotopically derived p_{CO_2} record shown in Fig. 2a is obtained. The Vostok ice-core p_{CO_2} record is shown for comparison. The p_{CO_2} records both demonstrate the sawtooth pattern typical of glacial–interglacial palaeochemical records²². Although misfits between the isotope and ice-core p_{CO_2} records are apparent and will be discussed below, the good correspondence strongly suggests that the isotope record contains information about CO_2 levels.

Assumptions about the equilibration between the oceans and the atmosphere can be decreased using a second approach. Data reported by McCabe⁴ and Rau *et al.*⁵ for mixed lacustrine and marine phytoplanktonic communities, respectively, yield slopes (a) of -17 and -21 . If we use $a = -20$ as an intermediate value (favouring the marine system), the range of CO_2 levels indicated by the isotope data exceeds that indicated by the ice core. This may be due to depletion of CO_2 in surface waters during the glacial period or to the presence of excess CO_2 in surface waters during the interglacial. The first alternative is consistent with suggestions that fertility of surface waters was enhanced during glacial intervals and with the observation that the net rate of accumulation of marine organic carbon during isotope stage 2 was 1.8 ± 1.0 times higher than that during isotope stage 1¹⁵. Accordingly, we can choose b to reflect an approach to equilibrium only during interglacial periods. If values of c are calculated from all observations of ϵ_P using $a = -20$ and $b = 2.5$, the dissolved CO_2 record shown in Fig. 2b is obtained. Also shown are the dissolved CO_2 concentrations that would be expected if surface ocean water were in equilibrium with Vostok CO_2 levels. Concentrations of dissolved CO_2 inferred during the last glacial period are low in comparison with those expected on the basis of ocean–atmosphere equilibration (for example, 5.3 compared

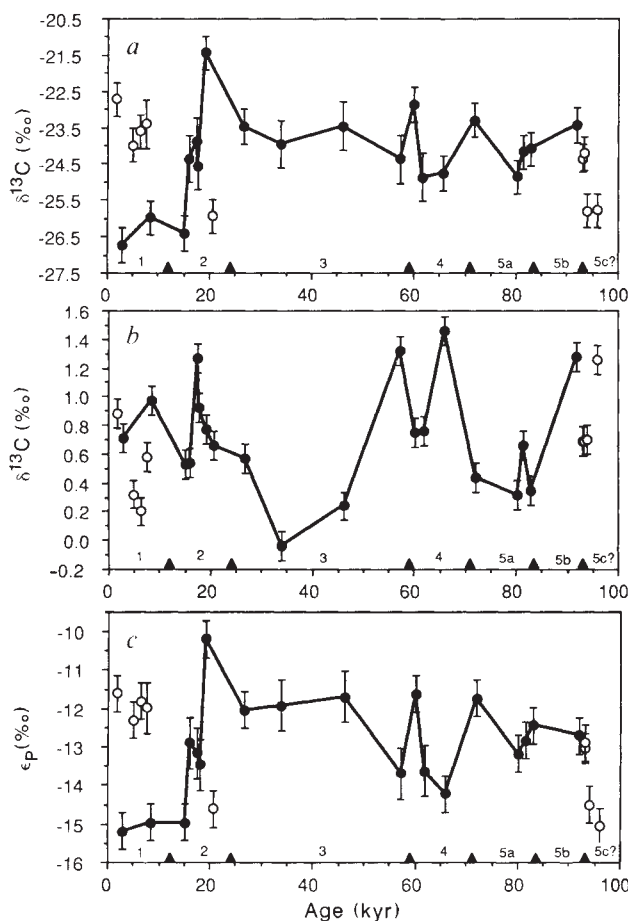


FIG. 1 a, Carbon isotopic compositions of samples of heptatriacontatriacontadien-2-one^{29,30} extracted from intervals of DSDP Hole 619¹⁰ (displaced sediments are represented by open circles). ($\delta^{13}C$ (‰) = $1,000[(R_{sample}/R_{PDB}) - 1]$, where $R = ^{13}C/^{12}C$ and $R_{PDB} = 0.0112372$). Ages were assigned from boundaries of the oxygen isotope stages¹⁴ and are compatible with three phytoplanktonic foraminiferal boundaries^{19,31}, a volcanic ash layer³², and an accelerator $\Delta^{14}C$ date (for details see ref. 15). b, Carbon isotopic compositions of calcite tests of *G. ruber*. c, Isotope fractionation between C_{37} alkadienones and *G. ruber* calcite expressed in terms of the fractionation factor ϵ_P (see equation (2)).

with $6.4 \mu\text{M}$ at 33.9 kyr, and 6.1 compared with $6.8 \mu\text{M}$ at 80.2 kyr). The indicated disequilibria correspond to Δp_{CO_2} values of 20–40 μatm , well within the range observed in modern oceans²¹.

Similar features appear in the sedimentary and ice-core records during the interval 50–90 kyr BP. Both, for example, indicate CO_2 maxima near 68 kyr BP, but these maxima and other points of similarity between the records are not temporally coincident. Given the oxygen isotope record in the sediment core¹⁴ and uncertainties regarding linear interpolation between stage boundaries¹⁵, it is possible, if not likely, that many of the misfits between the records reflect stratigraphic uncertainties.

The use of an isotope biomarker derived from a primary producer may be important in the recovery of the CO_2 -related signal. A p_{CO_2} record based on the isotope difference between dissolved CO_2 and total organic carbon (TOC) is shown in Fig. 2c. The curve is incongruent with the Vostok record of p_{CO_2} primarily because the isotopic composition of TOC has been strongly influenced by terrestrial contributions during the glacial intervals^{15,23}. This calculation illustrates the potential problem

of deriving p_{CO_2} values from sedimentary organic carbon of mixed origins. There are other cases in which the isotope relationship between marine primary material and TOC has been affected by secondary processes, and in which correct calibration of an ϵ - c relationship would be impossible without resolution of the primary material⁸. Selection of the alkenones as primary biomarkers instead of tetrapyrroles may have significant advantages. By focusing on a more restricted and uniform group of photosynthetic organisms, we may have avoided some ambiguities associated with varying biological responses to changes in p_{CO_2} . Coccolithophorids apparently lack a mechanism for accumulation of CO_2 (ref. 24), making them particularly suitable as recorders of CO_2 levels. Both coccoliths and their alkenone products are preserved in marine sediments, have existed since the Cretaceous²⁵, and are widely distributed geographically²⁶.

Our analyses of alkenones have allowed determination of the glacial-to-interglacial variation in the isotopic composition of the total biomass of the source organism³ and the recovery of a CO_2 -related signal covering ~ 100 kyr. This approach complements that of Broecker²⁷ and Shackleton *et al.*²⁸ who have shown that carbon isotope differences between benthic and planktonic foraminiferal carbonates can be interpreted in terms of total concentrations of dissolved inorganic carbon. Combining these and vertical nutrient distribution² approaches for the study of CO_2 in ancient oceans should allow exploration of the redistribution of carbon in the ocean and of mechanisms controlling atmospheric p_{CO_2} . The question unresolved here, whether low c values during the glacial interval accurately reflect depletion of CO_2 in the photic zone by carbon-fixing organisms, can be addressed by comparison of CO_2 records from oceans with varying histories of biological productivity. In the same way, areas of CO_2 uptake and release in ancient oceans may be detected. Derivation of a CO_2 -related record from marine sediments should be useful in developing a longer-term (from ~ 100 kyr to as much as 200 Myr) understanding of natural processes affecting atmospheric p_{CO_2} . □

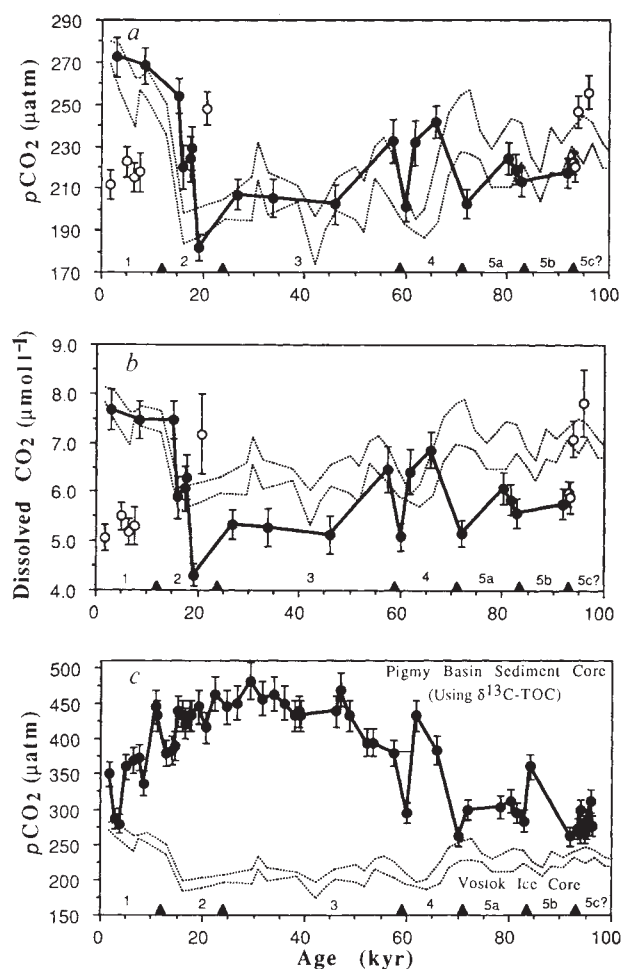


FIG. 2 a, Record of equilibrium atmospheric p_{CO_2} (see text) based on concentrations of dissolved CO_2 calculated using equation (1) with $a = -32.9$, $b = 14.3$ and ϵ_p values from Fig. 1c. The envelope (mean $\pm 2\sigma$) of the atmospheric CO_2 levels measured in the Vostok ice core¹ (dotted lines) is shown for comparison. b, Isotopically based record of sea surface concentrations of dissolved CO_2 (see text) compared with a hypothetical record (dotted lines) derived from the Vostok results based on assumption of ocean-atmosphere equilibration (see text). c, p_{CO_2} record based on assumption that $\delta_p = \delta_{\text{TOC}}$ (δ_{TOC} values from ref. 15). The Vostok ice core p_{CO_2} record¹ is shown for comparison.

Received 7 June; accepted 8 August 1990.

- Barnola, J. M., Raynaud, D., Korotkevich, Y. S. & Lorius, L. *Nature* **329**, 408–414 (1987).
- Boyle, E. A. *J. geophys. Res.* **93**, 15701–15714 (1988).
- Popp, B. N., Takigiku, R., Hayes, J. M., Landa, J. W. & Baker, E. W. *Am. J. Sci.* **289**, 436–454 (1989).
- McCabe, B. thesis, Univ. of Waikato (1985).
- Rau, G. H., Takahashi, T. & DesMarais, D. J. *Nature* **341**, 516–518 (1989).
- Fairbanks, R. G., Sverdrup, M., Free, R., Wiebe, P. H. & Be, A. W. H. *Nature* **298**, 841–844 (1982).
- Berger, W. H., Killingley, J. S. & Vincent, E. *Oceanologica Acta* **1**, 203–216 (1978).
- Hayes, J. M., Popp, B. N., Takigiku, R. & Johnson, R. *Geochim. cosmochim. Acta* **53**, 2961–2972 (1989).
- Freeman, K. H., Hayes, J. M., Trendel, J.-M. & Albrecht, P. *Nature* **343**, 254–256 (1990).
- Jasper, J. P. & Gagosian, R. B. *Paleoceanography* **4**, 603–614 (1989).
- Marlowe, I. T. *et al. Br. phycol. J.* **19**, 203–216 (1984).
- Bouma, A. H. *Am. Ass. Petrol. Geol. Mem.* **34**, 567–581 (1983).
- Bouma, A. H. *et al. Init. Rep. DSDP Leg 96*, 777–780 (1986).
- Williams, D. F. & Kohl, B. *Init. Rep. DSDP Leg 96*, 671–676 (1986).
- Jasper, J. P. & Gagosian, R. B. *Geochim. cosmochim. Acta* **54**, 1117–1132 (1990).
- Broecker, W. S. *Geochim. cosmochim. Acta* **46**, 1689–1705 (1982).
- Newman, J. W., Parker, P. L. & Behrens, E. W. *Geochim. cosmochim. Acta* **37**, 225–238 (1973).
- Ewing, M., Ericson, D. B. & Heezen, B. C. in *Habitat of Oil* (ed. Weeks, L. G.) 995–1054 (Am. Ass. Petrol. Geol., Tulsa, 1958).
- Kohl, B. *Init. Rep. DSDP Leg 96*, 657–670 (1986).
- Lorius, C. *et al. Nature* **316**, 591–596 (1985).
- Tans, P. P., Fung, I. Y. & Takahashi, T. *Science* **247**, 1431–1438 (1990).
- Broecker, W. S. & van Donk, J. *Rev. Geophys. Space Phys.* **8**, 169–189 (1978).
- Jasper, J. P. & Gagosian, R. B. *Nature* **342**, 60–62 (1989).
- Steeaman Nielsen, E. *Marine Photosynthesis* (Elsevier, Amsterdam, 1975).
- Farrimond, P., Eglinton, G. & Brassell, S. C. in *Advances in Organic Geochemistry 1985* (eds Leythaeuser, D. & Rullkötter, J.) 897–903 (Pergamon, Oxford, 1985).
- Okada, H. & McIntyre, A. *Micropaleontology* **23**, 1–13 (1977).
- Broecker, W. S. *Prog. Oceanogr.* **11**, 151–197 (1982).
- Shackleton, N. J., Hall, M. A., Line, J. & Carg, S. *Nature* **306**, 319–322 (1983).
- de Leeuw, J. W., van der Meer, F. W. & Rijpsma, W. I. C. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 211–217 (Pergamon, Oxford, 1980).
- Rechka, J. A. & Maxwell, J. R. *Tetrahedron Lett.* **29**, 2599–2600 (1988).
- Constans, R. E. & Parker, P. E. *Init. Rep. DSDP Leg 96*, 601–630 (1986).
- Ledbetter, M. T. *Init. Rep. DSDP Leg 96*, 685–688 (1986).

ACKNOWLEDGEMENTS. We thank the DSDP and ODP for providing samples for this research, L. P. Brzuz, K. H. Freeman and J. T. Hayes for mass spectrometric assistance, D. J. Hollander and K. H. Freeman for reviews of drafts of this report, and F. Prahl and K. Amthor for providing a culture of *Emiliana huxleyi* and the related alkenone extract. This work was supported by the NSF.

Uranium-mineralized micro-organisms associated with uraniferous hydrocarbons in southwest Scotland

A. E. Milodowski*, J. M. West*, J. M. Pearce*,
E. K. Hyslop†, I. R. Basham† & P. J. Hooker*

* British Geological Survey, Keyworth, Nottingham NG12 5GG, UK

† British Geological Survey, Murchison House, West Mains Road, Edinburgh EH9 3LA, UK

THE ability of microorganisms such as bacteria and fungi to concentrate uranium and other metals from solution has long been recognized (see, for example, refs 1–10), and has previously been applied to the economic recovery of metals from natural and industrial waste waters^{10–13}. This phenomenon is also important in the risk assessment of radioactive waste disposal, as the mobility of microbes may either enhance or retard radionuclide migration¹⁴. Although microbiological activity has been thought to influence the deposition or remobilization of uranium in natural deposits^{15,16}, there have been only a few direct observations of naturally mineralized microbial structures^{17,18}. During recent investigations of uranium mobilization from mineralized rocks in southwest Scotland, we observed the presence of uranium-mineralized structures attributable to the activity of filamentous microorganisms. Unlike previous accounts of either artificially stained^{1,3,6,19} or naturally mineralized microbes^{17,18}, the structures we describe display polymetallic mineralization, with a complex relationship between the metal species concentrated and its location in the microorganism.

Polymetallic mineralization containing uranium, copper, bismuth, cobalt and nickel (with arsenic and antimony) occurs at Needle's Eye (National Grid reference NX 915 562) on the north coast of the Solway Firth (Fig. 1). It lies largely in the contact zone of the late Caledonian Criffel granodiorite intrusive complex (397 ± 2 Myr²⁰) and hornfelsed Silurian metasedimentary rocks. These are faulted to the south against Carboniferous limestones and shales. Mineralization occurs only in a number of steeply inclined veins in northwest-trending fractures that cut the major coastal fault^{21,22} and ground waters flowing through fractures in the cliffs leach uranium (and other elements) from

it. Upwelling of ground water at the base of the cliffs together with uranium-charged surface run-off has redeposited uranium in the organic-rich Quaternary sediments (peat bog and flood plain silts) of the mudflats^{22,23}.

The uranium geochemistry of this site is being studied by the British Geological Survey to obtain data to validate the thermodynamic databases and equilibrium speciation codes used in modelling migration and retardation mechanisms of radionuclide transport pertinent to the disposal of radioactive waste²⁴. The primary mineralization and its alteration products have been reappraised using optical microscopy, secondary (SEM) and backscattered (BSEM) scanning electron microscopy with qualitative energy-dispersive X-ray microanalysis (EDXA), autoradiography and fission-track registration²². The distribution of radioactive phases was initially evaluated by registration of natural α -radioactivity in CR39 plastic²⁵. Uranium distribution and quantification was further refined by fission-track registration in Lexan polycarbonate plastic^{26,27}.

The characteristic primary mineralization comprises uraninite (var. pitchblende), brittle bituminous hydrocarbon, haematite, pyrite, chalcophyrite and native bismuth in a quartz-calcite-dolomite gangue. There are minor amounts of other ore minerals such as chalcocite, digenite, bornite, galena and clausthalite, and finely disseminated nickel-arsenic minerals. The paragenesis is complex and several episodes of mineralization are evident^{21,22}. Uraninite from the related nearby Steps Vein has been dated by the U–Pb method at 185 ± 20 Myr and the complete primary mineralization is believed to have occurred between that date and the time of the Lower Carboniferous²¹. Oxidative alteration near the surface has resulted in a wide variety of secondary uranium-silicate, uranium-copper-arsenic, calcium-arsenic, magnesium-arsenic and copper minerals^{21,22,28}.

Hydrocarbons from this locality have previously been reported to contain up to 35.89% by weight uranium^{29,30} and discrete inclusions of pitchblende^{21,29}. Fission-track registration techniques showed, however, that there was negligible intrinsic uranium within the hydrocarbons examined here but revealed high uranium levels (tens of per cent) on grain surfaces and in internal cracks (Fig. 2). BSEM-EDXA indicated that this was associated with fine-grained secondary precipitates of uranium silicate, uranium-calcium-arsenic and rarer uranium-magnesium-arsenic as coatings on, and replacing, the hydrocarbon substrate. Parnell²⁹ also found troegerite ($(\text{H}_3\text{O})_2(\text{UO}_2)$ -

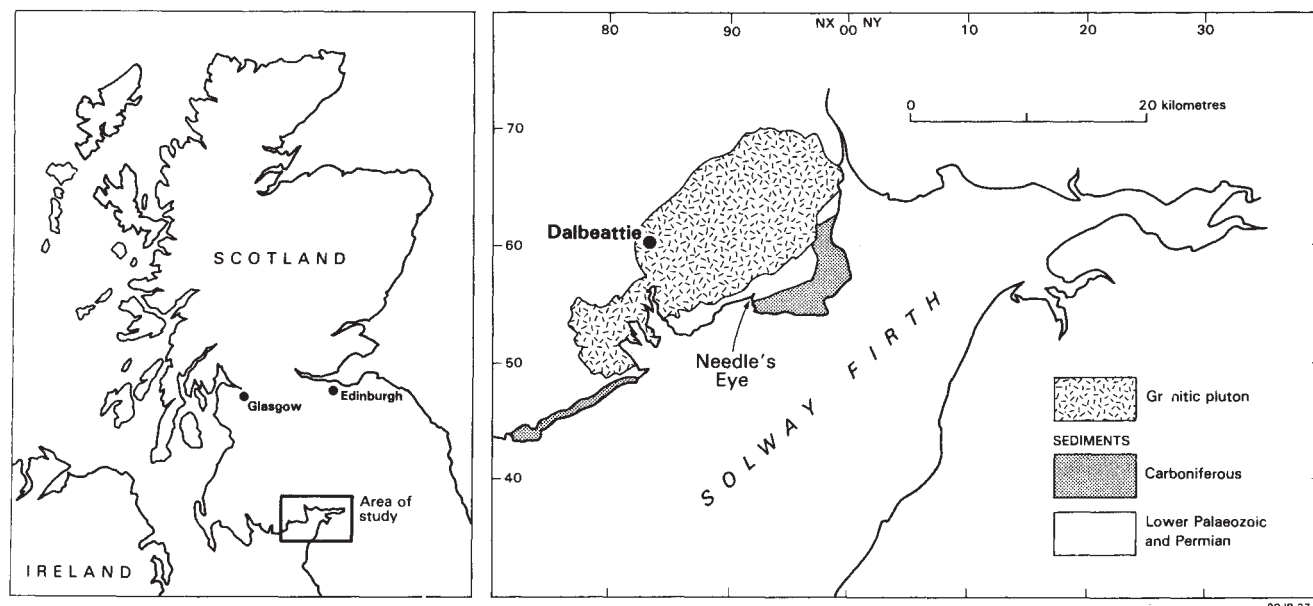


FIG. 1 Regional geology showing location of Needle's Eye site.

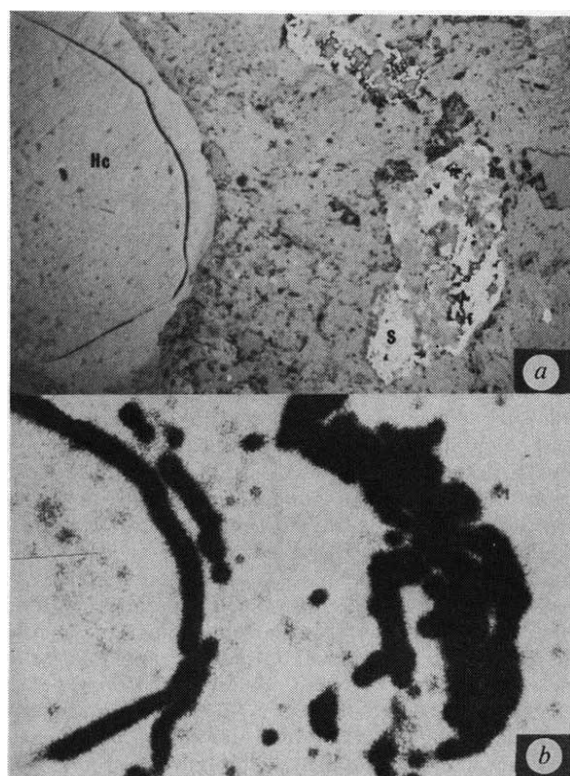


FIG. 2 *a*, Reflected-light micrograph of polished section of primary mineralization showing globular hydrocarbon (Hc) in a carbonate gangue containing sulphides (S). Field width, 5 mm. *b*, Fission-track print of *a* showing association of uranium with surfaces and cracks in the hydrocarbon, and with some sulphide minerals where chalcopyrite overgrowths have been replaced by uranium silicate.

$(\text{AsO}_4)_2 \cdot \text{H}_2\text{O}$) and zeunerite $(\text{Cu}(\text{UO}_2)_2(\text{AsO}_4)_2 \cdot 10\text{--}16\text{H}_2\text{O})$ as fracture coatings in the hydrocarbon globules. Uranium silicate is also closely associated with some of the sulphide minerals (Fig. 2) where it invariably replaces chalcopyrite overgrowths. In this context, the uranium seems to be texturally related to the mineral paragenesis of the primary hydrothermal vein²².

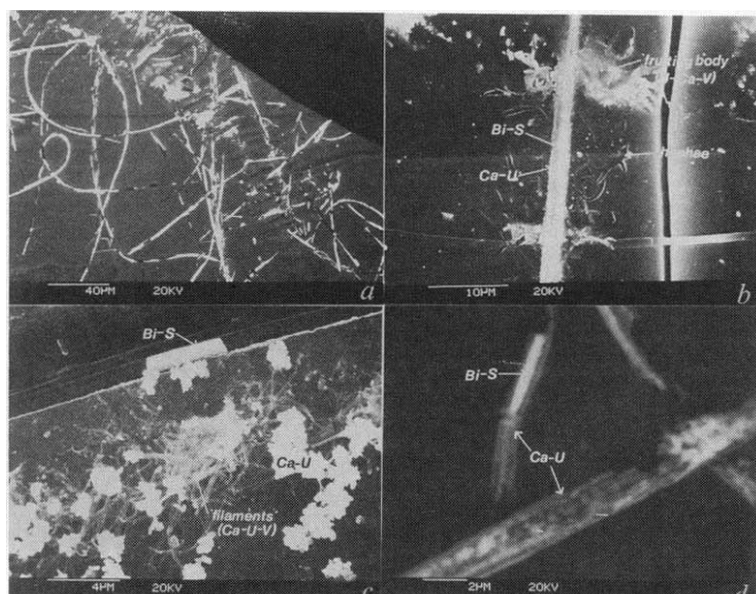
SEM-EDXA examination of the surfaces of open cracks in

hydrocarbon collected from an exposed outcrop and from loose mineralized blocks (dislodged from the cliffs during earlier uranium exploration²¹) lying in the peat bog revealed more complex associations of uranium and other metals. Peculiar filaments, apparently of biogenic origin and enriched in uranium and other metals, were observed to coat the crack surfaces (Fig. 3*a*). Some of the filaments are similar to fungal hyphae, but many possess a 'septate' structure and side-branching dendritic filamentous masses which more closely resemble the filaments and 'fruiting bodies' of actinomycetes³¹ (Fig. 3*b*). The main filaments are partially embedded in the surface of the hydrocarbon to a depth of $\sim 2\text{--}3\text{ }\mu\text{m}$, and have small unmineralized branching tubular structures which penetrate into the adjacent hydrocarbon (Fig. 3*b*). Finer masses of tangled filaments are also seen coating the fracture surfaces (Fig. 3*c*) but these are not necessarily related to the larger filaments. These finer filaments usually appear to be associated with an amorphous organic mucilage or biofilm.

There are distinct spatial relationships between each metal species and its location in these filaments. Dense bismuth or bismuth sulphide mineralization occurs in the 'core' regions of the larger filamentous bodies whereas the wall regions of the filaments are replaced, enriched or mineralized by a calcium-uranium mineral (Fig. 3*b, c*). Where cell structure can be recognized (Fig. 3*d*), a similar pattern is seen. Bismuth or bismuth sulphide mineralization occurs inside the cells and calcium-uranium enrichment occurs in the cell walls. In some cases selenium (apparently elemental) mineralization is seen in, or replacing the filament walls. The surfaces of dendritic 'fruiting bodies' and finer mucilaginous tangled filament masses are mineralized and enriched with calcium, uranium and vanadium. Aggregated masses of a fine-grained plate-like mineral composed of calcium and uranium (the tabular morphology possibly resembles that of liebigite³²) have formed in the biofilm associated with finer tangled filaments (Fig. 3*c*). All the metals identified in the biomineralized structures are most probably derived from the primary vein mineralization (such as uraninite, uranium silicate, native bismuth, clausenthalite).

The presence of viable microorganisms with similar morphology to the mineralized structures was demonstrated by simple inoculations of a variety of selective media (Czapek Dox agar, Tryptone soya agar and broth (Oxoid Ltd)) for yeasts, moulds and actinomycetes (Fig. 4). The metal enrichments observed were not necessarily the result of active accumulation by living microorganisms, however, and could have been the

FIG. 3 *a*, SEM micrograph showing morphology and distribution of uraniferous filaments on open fracture surfaces of vein hydrocarbon. *b*, SEM micrograph showing: the development of dendritic 'fruiting bodies' mineralized by uranium, calcium and vanadium; filament cores mineralized by bismuth or bismuth sulphide; filament walls mineralized by calcium and uranium. Unmineralized open channels penetrate the hydrocarbon adjacent to the main filament. *c*, SEM micrograph showing detail of fracture surface coated with mucilaginous tangle of fine organic filaments mineralized by calcium, uranium and vanadium, with nucleation of platelets of calcium-uranium mineral. *d*, BSEM micrograph showing cellular structure of filaments. The cell walls are selectively mineralized by calcium and uranium and the cell centre mineralized by bismuth or bismuth sulphide.



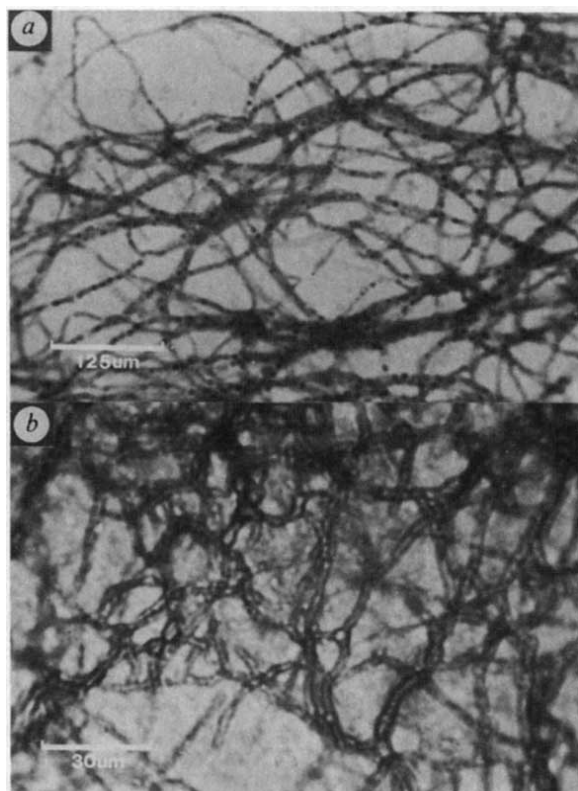


FIG. 4 Light microscopy micrographs of cultured microorganisms, stained with methylene blue. *a*, Morphology of filaments, scale bar 125 µm; *b*, light microscopy micrograph showing detail of branching filaments and possible spores, scale bar 30 µm. Oil immersion.

result of relatively recent surface-related processes. Experiments to demonstrate the potential for active uranium bioaccumulation on these microbes were outside the scope of this study. The high levels of toxic elements (uranium, selenium and bismuth) in these structures, however, indicate that post-mortem accumulation is more plausible¹³.

The association of uranium with cell walls of the microbes from Needle's Eye is similar to that described from laboratory experiments of uranium bioaccumulation^{1,3,6,19}. Site selectivity of metal uptake in bacteria is a recognized phenomenon related to ion-exchange or complexation between metal ions and active functional groups (such as amine, phosphate, phosphodiester and carboxyl groups) in polymers comprising the cell walls^{1,3,4,6-8,13}. The cell walls of many bacteria, including actinomycetes (which have a strong affinity for uranium⁹), are electronegative^{8,33,34} and are therefore able to act as ion-exchange resins^{7,8,13}, sequestering cations such as UO_2^{2+} from solution. Further uptake may result from inorganic adsorption or precipitation reactions¹³. Unlike previous accounts of either experimentally stained microorganisms or naturally mineralized microbes, the biomineralization in individual organisms from Needle's Eye is polymetallic. This is related to the complex geochemistry of the ground waters derived by leaching of polymetallic primary mineralization, in contrast with the simple chemistry of solutions used in experimental studies. Differences in the biogeochemical microenvironments associated with biofilms, cells walls and cell interiors may account for the different metal species deposited in these specific sites. □

Received 9 April; accepted 17 August 1990.

1. Beveridge, T. J. & Murray, R. G. E. *J. Bact.* **127**, 1502-1518 (1976).
2. Hawker, L. E. & Linton, A. H. *Microorganisms* (Arnold, London, 1979).
3. Beveridge, T. J. & Murray, R. G. E. *J. Bact.* **141**, 876-887 (1980).

4. Doyle, R. J., Matthews, T. H. & Streips, U. N. *J. Bact.* **143**, 471-480 (1980).
5. Disnar, J.-R., *Geochim. cosmochim. Acta* **45**, 363-379 (1981).
6. Beveridge, T. J., Meloche, J. D., Fyfe, W. S. & Murray, R. G. E. *Appl. env. Microbiol.* **45**, 1094-1108 (1983).
7. Ferris, F. G. & Beveridge, T. J. *Bioscience* **35**, 172-177 (1985).
8. Beveridge, T. J. *Biotechnol. Bioengng Symp.* **16**, 127-139 (1986).
9. Nakajima, A. & Sakaguchi, T. *Appl. Microbiol. Biotechnol.* **24**, 59-64 (1986).
10. Eccles, H. & Hunt, S. *Immobilisation of Ions by Bio-Sorption* (Ellis Horwood, London, 1986).
11. Horikoshi, T., Nakajima, A. & Sakaguchi, T. *J. Fermentation Technol.* **57**, 191-194 (1979).
12. Lorenz, M. G. & Krumbein, W. E. *Appl. Microbiol. Biotechnol.* **21**, 374-377 (1985).
13. Volesky, B. *Biotechnol. Bioengng Symp.* **16**, 121-126 (1986).
14. West, J. M., McKinley, I. G. & Chapman, N. A. *Radioactive Waste Management Nucl. Fuel Cycle* **3**, 1-15 (1982).
15. Magne, R., Berthelin, J. R. & Dommergues, Y. *Proc. Symp. Formation of Uranium Ore Deposits*, 73-88 (IAEA, Vienna, 1974).
16. Landais, P. *Terra Nova* **1**, 163-171 (1989).
17. Capus, G. & Munier, C. *C. r. heb. Séanc. Acad. Sci., Paris* **287**, 191-194 (1978).
18. Degens, E. T. & Venugopalan, I. *Nature* **298**, 262-264 (1982).
19. Beveridge, T. J. & Fyfe, W. S. *Can. J. Earth Sci.* **22**, 1893-1898 (1985).
20. Halliday, A. N., Stephens, W. E. & Harmon, R. S. *J. geol. Soc. Lond.* **137**, 329-348 (1980).
21. Miller, J. M. & Taylor, K. *Bull. Geol. Surv. G.B.* **25**, 1-18 (1966).
22. Basham, I. R., Milodowski, A. E., Hyslop, E. K. & Pearce, J. M. *Brit. geol. Surv. Fluid Processes Research Group Tech. Rep. WE/89/56* (Brit. geol. Surv., Keyworth, 1989).
23. Roberts, P. D., Ball, T. K., Hooker, P. J. & Milodowski, A. E. *Mater. Res. Soc. Symp. Proc.* **127** (eds Lutze, W. & Ewing, R. C.) 933-940 (MRS, Pittsburgh, 1989).
24. Côme, B. & Chapman, N. A. (eds) *Natural Analogue Working Group, Rep. CEC EUR 10315-EN* (CEC, Luxembourg, 1986).
25. Basham, I. R. *Econ. Geol.* **76**, 994-997 (1981).
26. Kleeman, J. D. & Lovering, J. F. *Atom. Energy Aust.* **10**, 3-8 (1967).
27. Basham, I. R., Ball, T. K., Beddoe-Stephens, B. & Michie, U. McL. *Proc. Symp. Uranium Exploration Methods* 385-397 (OECD/NEA, Paris, 1982).
28. Braithwaite, R. S. W. & Knight, J. R. *Miner. Mag.* **54**, 129-131 (1990).
29. Parnell, J. *Uranium* **4**, 197-218 (1988).
30. Eakin, P. A. *J. geol. Soc. Lond.* **146**, 663-673 (1989).
31. Buchanan, R. E. & Gibbons, N. E. (eds) *Bergey's Manual of Determinative Bacteriology*, 8th edn (Williams & Wilkins, Baltimore, 1975).
32. Palache, C., Berman, H. & Frondel, C. *Dana's System of Mineralogy*, Vol. 2 (Wiley, New York, 1951).
33. James, A. M. *Chem. Soc. Rev.* **8**, 389-418 (1979).
34. James, A. M. *Adv. Colloid Interface Sci.* **15**, 171-211 (1982).

ACKNOWLEDGEMENTS. We thank T. K. Ball for comments, E. J. Rowe for assistance, the Scottish Wildlife Trust for permission to work on the Southwick Coast Reserve, and the Scottish Universities Research Reactor Centre for the use of their irradiation facilities. The work has been supported by the UK Department of the Environment and the Commission of European Communities. The views expressed in this paper do not necessarily represent Government policy.

Infection of phytoplankton by viruses and reduction of primary productivity

Curtis A. Suttle, Amy M. Chan & Matthew T. Cottrell

Marine Science Institute, University of Texas at Austin, Port Aransas, Texas 78373-1267, USA

NATURAL marine waters contain roughly 10^6 to 10^9 virus particles per ml, yet their role in aquatic ecosystems and the organisms that they infect remain largely unknown. Electron microscopy has been used to study interactions between viruses and their hosts, focusing mainly on pathogens to prokaryotic organisms¹⁻⁵. Here we demonstrate that viral pathogens infect a variety of important marine primary producers, including diatoms, cryptophytes, prasinophytes and chroococcoid cyanobacteria. Also, addition to sea water of particles in the 0.002-0.2 µm size range, concentrated from sea water by ultrafiltration, reduced primary productivity (^{14}C bicarbonate incorporation) by as much as 78%. These results indicate that, in addition to grazing and nutrient limitation, infection by viruses could be a factor regulating phytoplankton community structure and primary productivity in the oceans.

Relatively large volumes of water (20-100 litres) were prefiltered through glass-fibre and membrane filters (142-mm diameter; 0.2-µm pore size) to remove zooplankton, phytoplankton and most bacteria. The remaining particulate matter in the filtrate was concentrated to 10-30 ml using Amicon hollow-fibre or spiral cartridge filters (molecular weight cutoff, 100,000 and 30,000, respectively). Negative-staining electron microscopy⁶ confirms that these methods yield diverse assemblages of virus-like particles essentially free of bacteria, clays and other particulates of non-viral appearance. We used the fluorochrome 4'-6-diamidino-2-phenylindole (DAPI) to

estimate the concentration of particulates in the concentrates that were associated with double-stranded DNA. Correcting for the concentrating factor of the cartridges, the number of DAPI-positive particles smaller than $0.2\ \mu\text{m}$ (presumed to be viruses) in the sea water ranged from 10^6 to $10^7\ \text{ml}^{-1}$ (Table 1). These concentrations are comparable to electron microscopy estimates²⁻⁵.

Isolation and screening of viruses infecting phytoplankton was achieved initially in liquid culture. Phytoplankton were grown exponentially (for several transfers) in batch culture using microwave-sterilized f/2-enriched⁷ natural sea water. To test for lytic viruses, growth rates and yields of unperturbed cultures were compared with cultures to which aliquots of natural viral communities, concentrated by ultrafiltration, had been added. Growth rates of the phytoplankton were monitored *in vivo* by chlorophyll fluorescence (an approximate measure of phytoplankton chlorophyll and hence biomass). When the addition of viral concentrate resulted in lower growth rates or cell lysis, the pathogen was purified by inoculating a small amount of supernatant from a centrifuged infected culture into an uninfected culture. The process was repeated many times to dilute non-replicating viruses from the original virus addition and to ensure that the effect could be propagated. The nature of the pathogen was confirmed by repeating these experiments using $0.2\ \mu\text{m}$ -filtered culture lysate as inoculum. Samples of culture lysate were also preserved for electron microscopy. Using this method we were successful in isolating several viruses that infect eukaryotic phytoplankton of diverse taxonomy, including a pennate and a centric diatom (both very small and of uncertain taxonomy), a cryptophyte (*Rhodomonas* sp.) and a prasinophyte (*Micromonas pusilla*). (We were unable to propagate the *Rhodomonas* virus, but lysis of the experimental cultures and the observation of viral-like particles in electron micrographs indicated that viruses probably were responsible for the treat-

ment effect.) In all cases addition of the viruses to cultures of their respective hosts caused dramatic decreases in fluorescence as the cultures approached stationary phase (Fig. 1). As far as we are aware, the only virus previously known to infect a marine eukaryotic phytoplankton was found in British Columbia coastal waters and, interestingly, was also a pathogen of *Micromonas pusilla*^{8,9}. Viruses have also been isolated that infect freshwater *Chlorella* spp. that are endosymbiotic in *Paramecium* and *Hydra*^{10,11}. We have also used ultrafiltration to isolate a phage that infects a chroococcoid cyanobacterium. As the alga and phage could withstand the temperatures required to pour top agar, this phage was further purified by combining centrifuged culture lysate with algae, embedding the mixture in a soft top layer of agar medium and screening for plaque formation¹².

As it was clear that particles in the 0.002 – $0.2\ \mu\text{m}$ size range (assuming a molecular weight of 30,000 corresponds to a particle diameter of about $0.002\ \mu\text{m}$) contained pathogens that infected a variety of phytoplankton, we investigated whether this fraction could affect rates of primary productivity. The 0.002 – $0.2\ \mu\text{m}$ fraction was concentrated from sea water and added back at a range of dilutions to 50-ml seawater samples from the same location. Samples were labelled with $5\ \mu\text{Ci}$ of [¹⁴C]bicarbonate and incubated under fluorescent lights at an irradiance of $120\ \mu\text{mol m}^{-2}\ \text{s}^{-1}$ for 4 h. Primary productivity, estimated from the radioactivity incorporated into particulate material, revealed that the rates of photosynthesis were as low as 22% of the controls without concentrate (Fig. 2). The inhibitor was also sensitive to autoclaving. We have repeated this experiment five times: on four occasions the results were similar to those

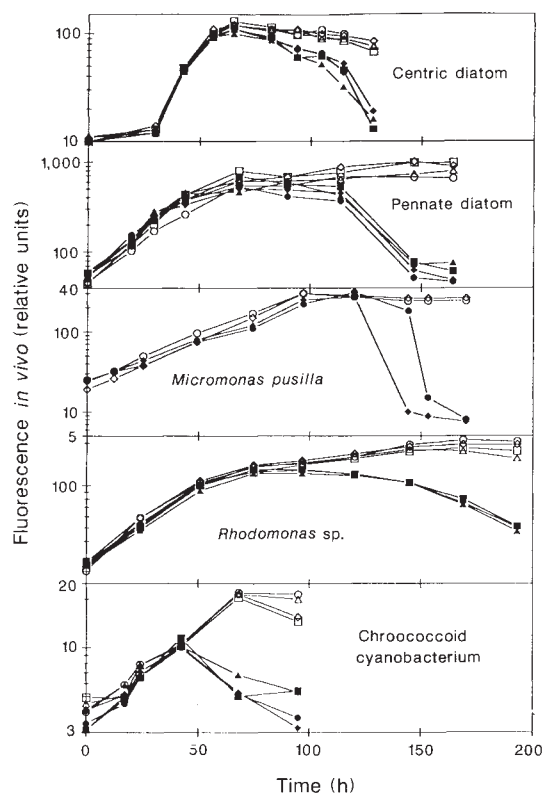


FIG. 1 Effect of adding naturally occurring marine viruses to cultures of their phytoplankton hosts. *In vivo* fluorescence gives a rough measure of chlorophyll *a* and hence of relative phytoplankton biomass. Each symbol represents a different replicate culture: open symbols, controls; closed symbols, cultures to which viruses have been added.

TABLE 1 Concentration of DAPI-positive particles in $0.2\text{-}\mu\text{m}$ filtered sea water

Location	Water type	DAPI-positive particles per ml	<i>n</i>
Laguna Madre	Hypersaline	3.3×10^7 (1.90×10^7)	6
MSI Pier/Harbour	Estuarine to clear coastal	1.4×10^7 (0.75×10^7)	6
Gulf of Mexico	Oligotrophic	2.0×10^6 (1.11×10^6)	3

Epifluorescence microscopy was used to obtain means and standard deviations of DAPI-positive particles in $0.2\text{-}\mu\text{m}$ filtered sea water in which the remaining particulate material was concentrated by ultrafiltration (*n*, number of sampling dates). The data are from several sampling dates between March and September 1989. Samples from Laguna Madre, a large ($1,660\ \text{km}^2$) hypersaline lagoon separated from the Gulf of Mexico by a barrier island were collected at $27^\circ 28.6'\ \text{N}$, $97^\circ 19.2'\ \text{W}$. The Marine Science Institute (MSI) Pier and Harbour are situated adjacent to Aransas Pass on the north end of a barrier island ($27^\circ 50.2'\ \text{N}$, $97^\circ 00.5'\ \text{W}$). Gulf of Mexico samples were obtained 70 km offshore from Aransas Pass. Estimates are from direct epifluorescent microscope counts of known volumes of DAPI-stained¹⁸ samples collected by ultrafiltration of sea water, which was prefiltered through glass-fibre and $0.2\text{-}\mu\text{m}$ membrane filters. Electron microscopy revealed that the concentrates were virtually free of bacteria or other particulate material that might interfere with the DAPI-counting method. Unstained samples did not contain fluorescent material. The DAPI-positive material was destroyed by heating to 95°C for 20 min. Samples were also treated for 30 min with protease-free DNase ($250\ \text{units ml}^{-1}$) to remove dissolved free DNA and DNA not protected by a protein coat. The method was tested by digesting a sample to which DNA had been added with DNase and monitoring the change in absorbance at 260 nm. The accuracy of this method in enumerating viruses was tested using known titres of two double-stranded DNA marine bacteriophages that we have isolated. Estimates of the efficiency of recovery were made by adding known numbers of marine bacteriophage to 20 litres of ultrafiltered (particle size $< 0.002\ \mu\text{m}$) sea water and subjecting it to our prefiltration and concentration procedures, then determining the concentration of DAPI-positive particles and plaque-forming units recovered. These results were compared with controls without bacteriophage. The overall concentration efficiencies of the spiral and hollow-fibre cartridges were 35 and 44% respectively.

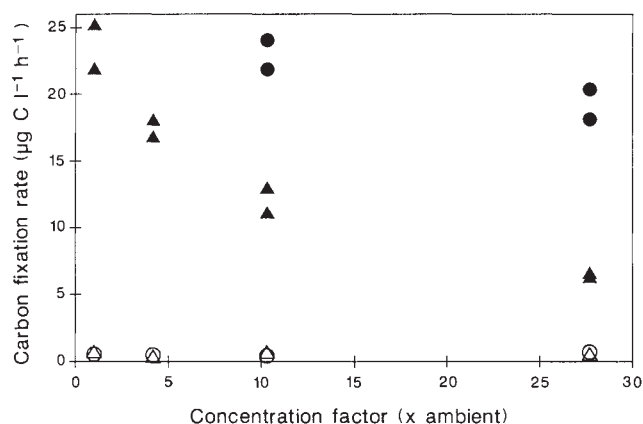


FIG. 2 Effect on phytoplankton photosynthetic rates of adding sea water from the 0.002–0.2 μm size fraction concentrated by ultrafiltration. Closed symbols represent samples that were incubated in the light and to which either autoclaved (circles) or non-autoclaved (triangles) concentrates were added. Open symbols represent samples that were incubated in the dark and to which either autoclaved (circles) or non-autoclaved (triangles) concentrates were added. On the basis of DAPI counts in the concentrate and the concentrating efficiency of the spiral cartridge, the ambient concentration of DAPI-positive particles in the size range 0.002–0.2 μm fraction was $2.1 \times 10^7 \text{ ml}^{-1}$.

described, with primary productivity being reduced to an average 44% (s.e. 21%), but in one instance there was no effect on photosynthesis when concentrate was added.

These data do not prove that viruses were responsible for the decrease in carbon-fixation rate: proteins of relatively high molecular weight would probably also be concentrated by our procedure, are heat labile and can have antibiotic activity. Whatever the causative agents, they were able to inhibit the principal primary producers in our system. If host-specific viruses were responsible, they evidently occur in close spatial and temporal proximity to the phytoplankton they infect.

It is significant that the phytoplankton for which we have isolated viral pathogens are representative of groups of algae that are important primary producers. Chroococcoid cyanobacteria are the most abundant primary producers in the oceanic environment and account for a substantial proportion of the total primary productivity in the nutrient-depleted waters that make up much of the world's oceans¹³. *Micromonas pusilla* is a small flagellate ($\sim 1 \mu\text{m}$ in diameter) which is found throughout the world in both coastal and oceanic habitats and which commonly exceeds 10^6 cells per litre¹⁴. Centric diatoms are the principal primary producers in upwelling and coastal areas, the most productive regions of the world's oceans. By contrast, pennate diatoms are major benthic primary producers and also constitute a significant component of marine fouling communities. *Rhodomonas* and other cryptomonads are a common element of coastal phytoplankton communities but remain poorly studied.

Virus-host isolates provide model systems by which to study the interactions of indigenous marine viruses with phytoplankton. Such systems will help us in determining the effect of viral infection on the primary productivity and structure of natural phytoplankton communities. As viruses can infect a variety of marine phytoplankton, they may affect pathways of nutrient and energy flow in the ocean, restrict or terminate phytoplankton blooms, maintain phytoplankton species diversity and serve as vectors for the transfer of genetic material between phytoplankton. The presence of algal viruses in sea water may also explain previous reports of lysis of natural phytoplankton communities¹⁵ and difficulties in balancing rates of primary productivity with loss rates for phytoplankton and sinks for organic carbon^{16,17}. □

Received 12 July; accepted 7 August 1990.

1. Sieburth, J. McN., Johnson, P. W. & Hargraves, P. E. *J. Phycol.* **18**, 416–425 (1988).
2. Bergh, Ø., Børsheim, G., Bratbak, G. & Heldal, M. *Nature* **340**, 467–468 (1989).
3. Proctor, L. M. & Fuhrman, J. A. *Nature* **343**, 60–62 (1989).
4. Børsheim, G. Y., Bratbak, G. & Heldal, M. *Appl. envir. Microbiol.* **56**, 352–356 (1990).
5. Bratbak, G., Heldal, M., Norland, S. & Thingstad, F. *Appl. envir. Microbiol.* **56**, 1400–1405 (1990).
6. Miller, S. E. *J. Electron Microsc. Tech.* **4**, 265–301 (1986).
7. Keller, M. D., Bellows, W. K. & Guillard, R. R. L. *J. exp. mar. Biol. Ecol.* **117**, 279–283 (1987).
8. Mayer, J. A. & Taylor, F. J. R. *Nature* **281**, 299–301 (1979).
9. Waters, R. E. & Chan, A. T. *J. gen. Virol.* **62**, 199–206 (1982).
10. Meintz, R. H., Schuster, A. M. & Van Etten, J. L. *Plant molec. Biol. Rep.* **3**, 180–187 (1985).
11. Van Etten, J. L., Xia, Y. & Meintz, R. H. in *Plant-Microbe Interactions* Vol. 2 (eds Kosuge, T. & Nester, E. W.) 307–325 (Macmillan, New York, 1987).
12. Van Etten, J. L., Burbank, D. E., Kuczmarski, D. & Meintz, R. H. *Science* **219**, 994–996 (1983).
13. Shapiro, L. P. & Guillard, R. R. L. in *Photosynthetic Picoplankton* (eds Platt, T. & Li, W. K. W.) *Can. Bull. Fish. aquat. Sci.* **214**, 371–389 (1986).
14. Thomsen, H. A. *Can. Bull. Fish. aquat. Sci.* **214**, 121–158 (1986).
15. Gieskes, W. W. C. & Elbrachter, M. *Neth. J. Sea Res.* **20**, 291–303 (1986).
16. Vargo, G. A. *et al. Limnol. Oceanogr.* **32**, 762–767 (1987).
17. Sellner, K. G. & Brownlee, D. C. in *Toxic Marine Phytoplankton* (eds Graneli, E., Sundstrom, B., Edler, L. & Anderson, D. M.) 221–226 (Elsevier, New York, 1990).
18. Coleman, A. W., Maguire, M. J. & Coleman, J. R. *J. Histochem. Cytochem.* **29**, 959–968 (1981).

ACKNOWLEDGEMENTS. We thank D. T. Brown, R. Mitchell and L. M. Proctor for help with electron microscopy and for discussion, and J. L. Van Etten and K. G. Sellner for their insight. J. A. Zeikus kindly provided phytoplankton from the University of Texas culture collection, and D. Stockwell obtained the sea water from Laguna Madre. This research was supported by grants from the US Office of Naval Research and from NOAA through the Texas A&M College Sea Grant program.

A terrestrial field experiment showing the impact of eliminating top predators on foliage damage

David A. Spiller & Thomas W. Schoener

Department of Zoology, University of California, Davis, California 95616, USA

CERTAIN ecological models^{1,2} predict that the impact of top predators on plants (producers) usually depends on the number of trophic levels in the food web: with three levels the impact is positive, whereas with four it is negative. These models assume that each of the nonbasal levels eats only the next level down. Indeed, freshwater pelagic systems show the predicted response when a fourth trophic level (piscivorous fish) is experimentally added to a three-level food web^{3,4}. In terrestrial systems, however, the top predator, if insectivorous, typically feeds on more than one trophic level⁵. This 'closed-loop omnivory'⁶, which may also characterize certain aquatic systems, necessitates a different model. *Anolis* lizards are often such top predators on small islands in the Bahamas⁷, eating both carnivorous and herbivorous arthropods^{8–10}. Numbers of web-spiders on islands without lizards are about 10 times those on islands with lizards^{11,12} and in experimental enclosures with lizards removed are 3 times those in enclosures with lizards at natural densities^{13,14}. Total leaf damage (by area) to buttonwood (*Conocarpus erectus*) on islands without lizards is 1.5 times that on islands with lizards¹⁵. These results indicate that lizards reduce the overall damage from herbivorous arthropods, even though they also reduce the numbers of web spiders which eat some herbivorous arthropods. Here we report on an experiment with lizards, spiders, herbivorous arthropods and sea grape, that indicates that the impact of top predators on producers depends on the relative strengths of interactions among the upper trophic levels. The impact is predicted to be positive if herbivorous arthropods are more prone to be eaten by lizards than by spiders (Fig. 1, model A) and negative if vice versa (model B). Experimental removal of lizards both increased scar damage on sea grape leaves, produced by homopterans and other organisms, and decreased damage produced by gall midges, indicating that the food web is a composite of models A and B (Fig. 1, model C). Contrary to the experiments in freshwater pelagic systems with four trophic levels, the net effect of top predators on producers was positive, apparently because the model A pathway predominates.

Details of the study site and experimental design are given elsewhere¹⁴. Briefly, in spring 1985 we staked out nine 83.6-m²

TABLE 1 Statistical analysis of leaf damage in the experimental plots 1985–88

	Lizard effect			Enclosure effect			Block effect		
	F	d.f.	P	F	d.f.	P	F	d.f.	P
MANOVA (Wilk's criterion)	56.38	3, 2	0.018	0.37	3, 2	0.786	9.36	6, 4	0.024
ANOVAs									
Scars	10.64	1, 4	0.031	0.07	1, 4	0.804	0.97	2, 4	0.453
Holes	0.28	1, 4	0.624	0.31	1, 4	0.607	0.42	2, 4	0.682
Galls	8.52	1, 4	0.043	0.01	1, 4	0.974	26.61	2, 4	0.005

Data are in Fig. 2. A repeated-measures two-way MANOVA was performed on mean cumulative leaf damage (percentage of entire leaf area) in each plot at the end of each year; the dependent variables were scars, holes and galls, and the main effects were treatments and blocks. The orthogonal decomposition method²⁷ was used to test *a priori* hypotheses on the treatment effect (as in ref. 14): the 'lizard effect' was determined by contrasting lizard-removal enclosures with (control enclosures with lizards plus unenclosed plots with lizards), and the 'enclosure effect' was determined by contrasting control enclosures with unenclosed plots. The 'block effect' compares plots with low, medium and high vegetation heights. Separate univariate tests (repeated measures two-way ANOVAs) were performed on each dependent variable, using the convention that when an overall main effect is significant at the 0.05 level in the multivariate test, the *P* values in the univariate tests can be used to evaluate the main effect on each variable²⁸. Percentage damage was arcsin-square-root-transformed.

plots on Staniel Cay. We stratified the plots into three blocks on the basis of mean vegetation height (low, medium, high) with three plots per block. In each block, enclosures were built on two plots and the third plot was left unenclosed; one enclosure was randomly chosen from which to remove lizards and the other enclosure was the control with lizards present at their natural density. Thus, there were three treatments in each of the three blocks—lizard-removal enclosures, control enclosures with lizards and unenclosed plots with lizards. Comparisons between control enclosures and unenclosed plots test for the effect of enclosures *per se*. The enclosures consisted of wood-framed fences, with hardware cloth ($\frac{1}{8}$ -inch mesh) attached to the sides. To prevent lizards from entering or leaving the enclosures, a continuous 0.4-m-wide strip of polypropylene plastic was mounted on top of each fence, forming 0.2-m overhangs on the inside and outside of each enclosure; except for the overhangs, the enclosures were open on top.

We measured herbivory on sea grape (*Coccoloba uvifera* L.), the dominant plant at the study site. Its leaves are large, coriaceous and roughly elliptical. From 1985–88 we measured the amount of damage that accumulated on a large sample of new leaves in each plot; individual leaves were tagged and monitored from May to late autumn. Damaged areas were classified as scars, holes or galls. Areas that appeared to be damaged by physical factors (for example, burning) were excluded. Scars are well defined, necrotic areas, and holes are entirely missing areas. In the field we observed scar damage produced by homopterans (leafhoppers and aphids) and an undetermined miner. We observed hole damage produced by a coleopteran (June beetle), lepidopteran larvae and hymenopterans (a leafcutter ant and a leafcutter bee). Some scar and hole damage might have been caused by microorganisms. All galls were produced by a cecidiomyiid midge (probably *Ctenodactylomyia watsoni*^{16,17}, identified by R. L. Gagné).

Scar damage was the most abundant type of herbivory (Fig. 2). In each year, scar damage tended to be higher in lizard-removal enclosures than in control enclosures with lizards; control enclosures and unenclosed plots with lizards had about the same amount of scar damage. Overall ('mean yearly total' in Fig. 2) scar damage in lizard-removal enclosures was 2.2 times that in control enclosures, and was nearly the same in control enclosures and unenclosed plots. Hole damage was relatively low from 1985 to 1987, and the treatments were about the same. In 1988 hole damage increased and tended to be higher in lizard-removal enclosures than in treatments with lizards. Overall hole damage was 1.7 times higher in lizard-removal than in control enclosures, and about the same in control enclosures and unenclosed plots. In contrast to scars and holes, gall damage in 1985 tended to be lower in lizard-removal enclosures than in treatments with lizards. Although patterns were not consistent in each year, overall gall damage was 1.7 times higher in control than in lizard-removal enclosures, and was nearly the same in control enclosures and unenclosed plots.

The three kinds of leaf damage taken together responded significantly to the removal of lizards (multivariate analysis of variance (MANOVA) in Table 1). Individual tests (analysis of variance (ANOVA) in Table 1) showed that lizards had a significant effect on scar and gall damage but not hole damage. No enclosure effect was significant. The block effect on gall damage was significant; pairwise comparisons of high, medium and low vegetation blocks showed that for gall damage, high > medium, high > low and medium the same as low (Tukey's multiple comparison procedure, *P* < 0.05). The block effect was not significant for scars or holes.

The experiment revealed that lizards had a negative effect on scar damage, a nonsignificant effect on hole damage and a positive effect on gall damage. We suggest that the negative effect on scar damage was caused, at least in part, by lizards

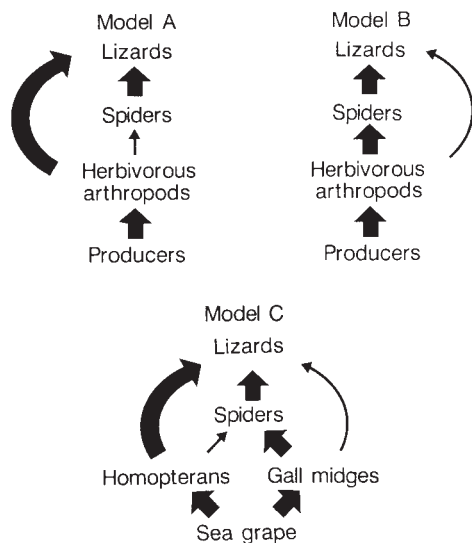
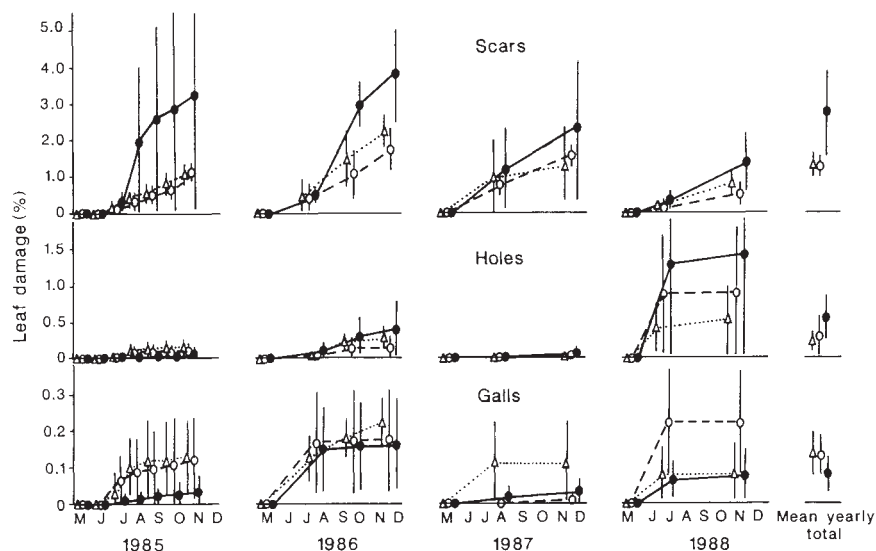


FIG. 1 Alternative models for (partial) food webs on Bahamian subtropical islands. Model A, model dominated by lizard predation on herbivorous arthropods. Model B, model dominated by spider predation on herbivorous arthropods. Model C, composite model (having both A and B portions) supported in this study. Arrow thicknesses are interpreted as follows. An arrow from *x* to *y* gives the effect of *y* on *x* in the absence of effects on *y* from other elements of the food web that connect to *y*, that is, predators on *y* or alternative prey of *y*. In particular, the population size of *y* as directly affected by other elements of the web is not reflected in arrow thickness. Arrow thickness, among other things (see ref. 5 for elaboration), can be interpreted as a per capita effect of *y* on *x*.

FIG. 2 Mean cumulative scar, hole and gall damage per leaf (percentage of the entire leaf area) on new sea grape leaves from May to late autumn 1985, 1986, 1987 and 1988 in enclosures with lizards removed (●—●), control enclosures with lizards at natural densities (○—○) and unenclosed plots with lizards at natural densities (△—△). Each treatment had three replicate plots. Each point is the mean of the mean cumulative leaf damage in each plot of a treatment. 'Mean yearly total' is the mean of mean cumulative leaf damage at the end of each year in a given plot. Error bars are ± 1 s.d. of the mean of the plot means for each treatment.

METHODS. In May 1985 numbered tags were put on the petioles of 12 new, undamaged leaves in each plot. The distribution of leaves in each plot consisted of four uniformly spaced patches with three leaves per patch. At about monthly intervals, for each leaf and each damaged area on the top side, the maximum diameter and the diameter perpendicular to the maximum were measured *in situ* with a ruler. The same procedure was repeated in 1986, 1987 and 1988, except 15 leaves were tagged in each plot (arranged in five uniformly spaced patches with three leaves per patch), and measurements were taken at about three-monthly intervals. Thus in the entire study, 513 leaves were repeatedly measured. Leaf areas and damaged areas were estimated by assuming that they were elliptical. Percentage damage per leaf ($100 \times \text{sum of damaged}$



areas/entire leaf area) was computed separately for scars, holes and galls. Most of the tagged leaves died during late autumn and winter. For each leaf that abscised before the final census of the year, the percentage damage when last measured was included in the subsequent calculations of mean cumulative leaf damage in the plot. M-D, May to December.

eating leafhoppers and aphids. At the study site, these insects were common in lizard diets¹⁸. On the other hand, the insects that we observed producing holes on sea grape leaves were sometimes larger than the gape of most *Anolis* lizards. This may explain why the effect of lizards on hole damage was not significant.

The positive effect of lizards on gall damage is intriguing. Small Diptera (such as gall midges) were rare in lizard diets but very common in spider webs (ref. 18, and our unpublished data). Numbers of web-spiders in lizard-removal enclosures were three times those in treatments with lizards¹⁴. Hence, we suggest that reduced gall damage in lizard-removal enclosures was caused by increased predation of gall midges by spiders. Other evidence for this kind of effect comes from an experiment in the Lesser Antilles by Pacala and Roughgarden¹⁹. They found that lizards reduced the abundances of web spiders and ground insects and increased the abundance of aerial insects; they suggested the former resulted from direct predation, whereas the latter resulted indirectly, from decreased predation by web-spiders.

The effect of top predators on producers is an example of an indirect effect, one that is mediated by one or more intervening trophic elements. Indirect effects might be expected to be weaker than direct effects, or at least to become established more slowly after perturbation²⁰⁻²³. This study supports the first expectation, in that the (mainly) direct effect of lizards on spiders was threefold, whereas the indirect effects were closer to twofold. Within the time course of the experiment (Fig. 2), however, no evidence arises for the second expectation—effects of herbivory were evident within a year of lizard removal, and four years of observation were necessary mainly to obtain stable averages for each plot, as herbivory was sometimes very spotty in any one year. Moreover, especially in food webs where indirect effects are possible, our study shows that experiments are effective in revealing the predominant pathways, pathways unspecified by connection patterns alone.

Our experimental results can be accounted for by a composite of two previously proposed models (Fig. 1, model C); scar-producing arthropods, for example, homopterans, are directly affected by lizards more than spiders, whereas gall-producing arthropods, gall midges, are directly affected by spiders more than by lizards. Because scars accounted for most of the damage,

lizards reduced total herbivory. This is consistent with a comparative study of buttonwood leaf damage on small islands¹⁵. Hence, for total herbivory the pathway in model A is predominant. In model A the effect of the top level on herbivores and producers is in agreement with the three-level scheme of Hairston *et al.*²⁴ proposed for terrestrial systems. Certain experiments in freshwater pelagic systems with four trophic levels³, however, show the opposite effect, favouring the equivalent of model B (freshwater webs may often have three or fewer levels as well⁴, and removal of the top predator in such webs may or may not affect lower levels²⁵). A major difference between our terrestrial system and such four-level freshwater pelagic systems is that the size difference between top predators and herbivores is greater in the latter (adult piscivorous fish versus herbivorous zooplankton) than in the former (lizards versus herbivorous arthropods). Consequently, except during early juvenile stages, piscivorous fish do not eat herbivores, and the interaction between top predators and herbivores is weak, as in model B. Indeed, in freshwater systems showing less size difference between top predators and herbivores, model A may hold. Thus, as discussed elsewhere^{5,21,23,26}, body-size relationships between predators and prey may weigh heavily in determining the main pathways in food webs. □

Received 3 April; accepted 13 July 1990.

1. Fretwell, S. D. *Persp. Biol. Med.* **20**, 169-185 (1977).
2. Oksanen, L., Fretwell, S. D., Arruda, J. & Niemala, P. *Am. Nat.* **118**, 240-261 (1981).
3. Carpenter, S. R. & Kitchell, J. F. *BioScience* **38**, 764-769 (1988).
4. Persson, L., Andersson, G., Hamrin, S. F. & Johansson, L. in *Complex Interactions in Lake Communities* (ed. Carpenter, S. R.) 45-65 (Springer, New York, 1988).
5. Schoener, T. W. *Ecology* **70**, 1559-1589 (1989).
6. Pimm, S. L. & Lawton, J. H. *Nature* **275**, 542-544 (1978).
7. Schoener, T. W. & Schoener, A. J. *Anim. Ecol.* **52**, 237-262 (1983).
8. Schoener, T. W. *Ecology* **49**, 704-726 (1968).
9. Schoener, T. W. & Gorman, G. C. *Ecology* **49**, 819-830 (1968).
10. Floyd, H. B. & Jenssen, T. A. *Copeia* **1983**, 319-331 (1983).
11. Schoener, T. W. & Toft, C. A. *Science* **219**, 1353-1355 (1983).
12. Toft, C. A. & Schoener, T. W. *Oikos* **41**, 411-426 (1983).
13. Schoener, T. W. & Spiller, D. A. *Science* **236**, 949-952 (1987).
14. Spiller, D. A. & Schoener, T. W. *Ecol. Monogr.* **58**, 51-77 (1988).
15. Schoener, T. W. *Oikos* **53**, 253-266 (1988).
16. Gagné, R. L. *The Plant-feeding Gall Midges of North America* (Cornell Univ. Press, Ithaca, 1989).
17. Johnson, W. T. & Lyon, H. H. *Insects that Feed on Trees and Shrubs* (Cornell Univ. Press, Ithaca, 1988).
18. Spiller, D. A. & Schoener, T. W. *Oecologia* **83**, 150-161 (1990).
19. Pacala, S. & Roughgarden, J. *Oecologia* **64**, 160-162 (1984).
20. Lawton, J. H. & Pimm, S. L. *Nature* **272**, 189-190 (1978).

21. Pimm, S. L. *Food Webs* (Chapman & Hall, London, 1982).
22. Brown, J. H., Davidson, D. W., Munger, J. C. & Inouye, R. S. in *Community Ecology* (ed. Diamond, J. M. & Case, T.) 41–61 (Harper & Row, New York, 1986).
23. Lawton, J. H. in *Ecological Concepts* (ed. Cherrett, J. M.) 43–78 (Blackwell Scientific, Oxford, 1989).
24. Hairston, N. G., Smith, F. E. & Slobodkin, L. B. *Am. Nat.* **94**, 421–425 (1960).
25. Leibold, M. A. *Am. Nat.* **134**, 922–949 (1989).
26. Stein, R. A. et al. in *Complex Interactions in Lake Communities* (ed. Carpenter, S. L.) 161–179 (Springer-Verlag, New York, 1988).
27. Sokal, R. R. & Rohlf, F. J. *Biometry* (Freeman, San Francisco, 1981).
28. Stevens, J. *Applied Multivariate Statistics for the Social Sciences* (Erlbaum, Hillsdale, New Jersey, 1986).

ACKNOWLEDGEMENTS. We thank S. Pimm and M. Power for comments, and H. Hartman and G. Gustavson for permission to use their land. Supported by the NSF.

Oxygen consumption of bumblebees in forward flight

C. P. Ellington*, K. E. Machin*† & Timothy M. Casey‡

* Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

‡ Department of Entomology, New Jersey Agricultural Experiment Station, Cook College, Rutgers University, New Brunswick, New Jersey 08903, USA

AERODYNAMIC theories predict that the mechanical power required for flapping animal flight varies with flight speed according to a 'U-shaped' curve^{1–6}. At some intermediate speed there is a pronounced power minimum, typically half the power required for hovering flight. These theories are widely used for birds and bats, but the predictions are not well supported by measurements of the rate of energy expenditure, or metabolic rate—usually determined from the rate of oxygen consumption—of animals flying in wind tunnels. This rate follows a distinctly U-shaped curve in a few species, but quantitative agreement with theoretical predictions is nevertheless unsatisfactory^{7–9}. In other species either the curve is much flattened or the rate shows no significant change with flight speed^{10–13}. In one case the agreement between theory and experiment is remarkable over the limited speed range under test^{6,14}, but there is a discrepancy in the shapes of the two curves which would lead to large differences outside that speed range. Hummingbirds show no significant difference in their flight metabolism over speeds from hovering up to 7 m s⁻¹ (ref. 15). At even higher speeds the metabolic rate increases, presumably as a result of the increased power needed to overcome body drag. The theories can also be applied to insects, but there are no equivalent data against which they can be tested. Here we report such measurements for bumblebees in free, controlled, forward flight over a speed range of 0–4 m s⁻¹. The metabolic rate does not vary significantly with flight speed, suggesting that the theories are inadequate for insect flight.

Like hummingbirds, insects are ideal candidates for a test of the theories because most are capable of flight over a wide speed range, including hovering. The metabolic rate of insects in free forward flight has not been reported previously, although a few values have been published for tethered insects at one or two flight speeds^{16,17}. A small mask is normally used to collect expired gases from vertebrates for analysis, but the mask required for the spiracle system of insects¹⁸ is too cumbersome for free flight. The alternative approach of measuring the decline in oxygen levels in a closed-circuit wind tunnel is beyond the capabilities of unmodified commercial instruments. In tests of the Ametek S3A/II differential oxygen analyser we demonstrated that its resolution and accuracy are mainly limited by amplifier noise and drift. Its performance was therefore improved by using a custom-built nanovolt amplifier instead. Better control over the temperature of the heated sensor (750 °C) was also achieved by stabilizing the power supplies of the heater circuit.

This modified analyser monitored the difference in partial pressure of oxygen between a variable speed (0–4 m s⁻¹) closed-

circuit wind tunnel and a reference chamber of similar volume (4 litres). Just before entry into the sensor, each gas circuit was connected to a partially inflated condom that would change volume in response to slight pressure changes. This closed-circuit, constant-pressure, variable-volume differential respirometry system greatly attenuated pressure disturbances due to the tunnel fan, the gas sampling pumps and temperature variations in the closed circuits. A qualitative ranking of bumblebee activity levels was displayed on a chart recorder simultaneously with the oxygen measurement.

Individual bumblebees were introduced into the tunnel through a small porthole, which nevertheless allowed some mixing with room air. Bumblebees were therefore kept in the dark for a few minutes to minimize their activity while the oxygen level equilibrated, and while noise and drift were checked on the chart recording. The lighting was then switched on to stimulate activity. A sample flight experiment is shown in Fig. 1; note that when the bumblebee is in continuous, oriented flight the oxygen level drops linearly. The rate of oxygen consumption was calculated from the slope only during such periods, and corrected to standard temperature and pressure, dry. Lights were turned off after the flight, causing the metabolic rate to drop while the recording was checked for drift again. Finally, a volume of pure O₂ was injected into the tunnel gas circuit; this calibration injection permitted direct conversion of recorder deflections to units of ml O₂.

The rate of oxygen consumption for bumblebees flying at speeds up to 4 m s⁻¹ was generally independent of speed, and comparable with values obtained during hovering flight inside the reference chamber (Fig. 2). The maximum speed of the wind tunnel (4 m s⁻¹) is close to the maximum recorded flight speed (5 m s⁻¹) for bumblebees¹⁹. Three flights at each speed were obtained for one of the bees, revealing that the variability in flight cost is similar over the speed range. A shallow U-shaped curve is suggested by some of the results, but this small curvature usually lies within the experimental errors.

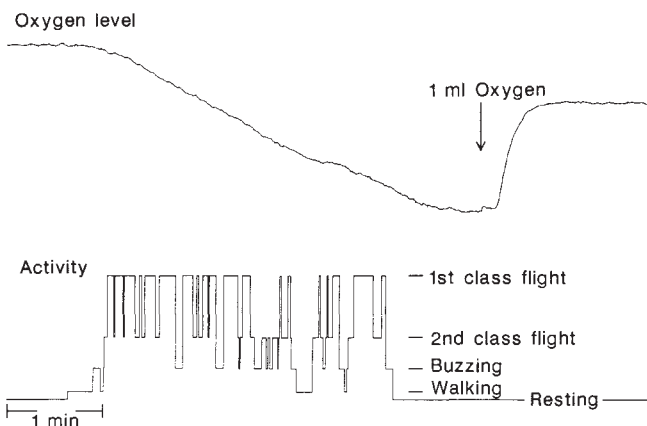


FIG. 1. Sample chart recording showing the drop in oxygen level in the wind tunnel during flight at 1 m s⁻¹ of a *Bombus lucorum* queen weighing 464 mg, followed by a calibration injection of 1 ml O₂ into the tunnel gas circuit. Activity levels are displayed on the lower trace: first-class flight is in the centre of the working section with good orientation; second-class flight is either close to a wall or with poor orientation; buzzing indicates attempted flight while a leg was caught in the filter of the tunnel; walking and resting are self-explanatory. Orientation of the bumblebees was controlled using a striped drum rotating around an 8 W ultraviolet fluorescent lamp above the front of the working section (10 cm square by 17 cm long). Tunnel and reference chamber had the same initial gas composition (dry air, free from CO₂ at 20.9% O₂), and were temperature controlled to 32 ± 0.05 °C. Gas samples were pumped continuously at 800 ml min⁻¹ from the chambers, passed through CO₂ and H₂O vapour scrubbers (soda lime and silica gel, respectively), run through the analyser, and then returned to the chambers. Changes in O₂ level were calibrated against injections of pure O₂ into the circuits. Accuracy and sensitivity were 0.001% and 0.0003% O₂, respectively; three times better than the S3A/II operating under the most favourable conditions with very low gas flow rates.

† Deceased.

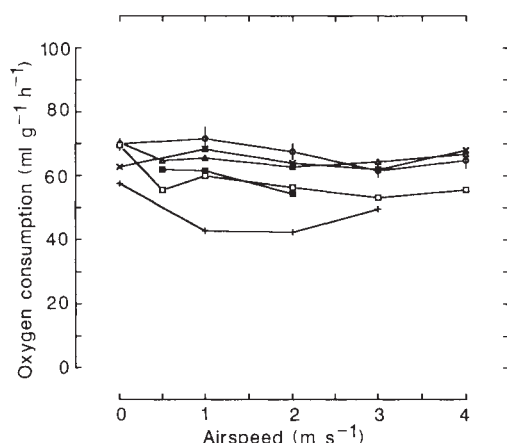


FIG. 2 Oxygen consumption as function of airspeed for six bumblebees. Symbol identification and mean mass (mg) during experiments are as follows: three *Bombus lucorum* queens (■, 537; ○, 533; □, 464); one *B. lucorum* worker (△, 236); two *B. pascuorum* workers (+, 301; ×, 181). Error bars for the 533 mg queen are for 2 standard errors.

The speed corresponding to the minimum mechanical power requirement can be calculated from the aerodynamic theories, giving values of about $3\text{--}4\text{ m s}^{-1}$ for the bumblebees in Fig. 2. Although these speeds are not inconsistent with the experimental data, a large discrepancy exists in the magnitude of the power estimates; according to theory, the power for hovering is about twice the minimum power, but this pronounced difference is absent in the measurements. Furthermore, a detailed aerodynamic analysis of bumblebee flight using full kinematic data on the wing motions has recently predicted mechanical power requirements with a very shallow U-shaped curve²⁰, closely matching the experimental results. The flight costs calculated from the usual theories can thus be in error by up to a factor of two over the range of flight speeds. A similar comparison of the theories with metabolic data and detailed aerodynamic analyses is sorely needed for vertebrate flight. Until then, these studies on bumblebees provide the most comprehensive test, and the shortcomings of the theories are clearly evident.

The attractiveness of the more popular theories lies in their simplicity; from a few morphological measurements the entire power curve can be constructed. It would not be surprising, however, if the accuracy of these simple analyses was limited. Even Pennycuik, the main advocate of such theories, has always advised caution in using them outside the normal range of cruising speeds. Nevertheless, the theories are often used indiscriminately for a variety of predictions and conclusions about energetics and flight performance. Given the discrepancies now found for bumblebees as well as those for vertebrate fliers, we recommend even greater caution in using these theories until such discrepancies have been resolved. For insects and hummingbirds, at least, we suggest an interim method for estimating their flight costs. Because flight metabolism of the bumblebees and hummingbirds is essentially constant over most of their normal range of speeds, investigators could simply use values obtained in free hovering flight for the species of interest. Such data are available for hummingbirds^{21,22} and a wide variety of insects²³, and these values should be much closer to the metabolic cost of forward flight than any simple power curve prediction. The only factor known to cause substantial variation of metabolic rate in individual insects during free flight is the payload they carry^{24,25}, and this is easily quantified by weighing the animal soon after capture. □

8. Thomas, S. P. *J. exp. Biol.* **63**, 273–293 (1975).
9. Carpenter, R. E. *J. exp. Biol.* **114**, 619–647 (1985).
10. Tucker, V. A. *Am. J. Physiol.* **222**, 237–245 (1972).
11. Bernstein, M. H., Thomas, S. P. & Schmidt Neilsen, K. *J. exp. Biol.* **58**, 401–410 (1973).
12. Torre-Bueno, J. R. & LaRoche, J. *J. exp. Biol.* **75**, 223–229 (1977).
13. Hudson, D. M. & Bernstein, M. H. *J. exp. Biol.* **103**, 121–130 (1983).
14. Rothe, H.-J., Biesel, W. & Nachtigall, W. *J. comp. Physiol.* **B 157**, 99–109 (1987).
15. Berger, M. in *BIONA Report 3* (ed. Nachtigall, W.) 307–314 (G. Fischer, Mainz, 1985).
16. Kammer, A. E. & Heinrich, B. *Adv. Insect Physiol.* **13**, 133–228 (1978).
17. Nachtigall, W., Rothe, U., Feller, P. & Jungmann, R. *J. comp. Physiol.* **B 158**, 729–737 (1989).
18. Armstrong, G. & Mordue, W. *Physiol. Ent.* **10**, 353–358 (1985).
19. Demoll, R. *Der Flug der Insekten und der Vögel* (Fischer, Jena, 1918).
20. Dudley, R. & Ellington, C. P. *J. exp. Biol.* **148**, 53–88 (1990).
21. Berger, M. & Hart, J. S. *J. comp. Physiol.* **81**, 363–380 (1972).
22. Berger, M. *J. Orn., Lpz.* **115**, 273–288 (1974).
23. Casey, T. M. in *Insect Flight* (eds Goldworthy, G. & Wheeler, C.) 257–272 (CRC, Boca Raton, 1989).
24. Heinrich, B. *J. comp. Physiol.* **96**, 155–166 (1975).
25. Wolf, Th. J., Schmid-Hempel, P., Ellington, C. P. & Stevenson, R. D. *Funct. Ecol.* **3**, 417–424 (1989).

ACKNOWLEDGEMENTS. We thank A. Bloom for use of the nanovolt amplifier and for advice on the oxygen analyser. This work was supported by the SERC and NSF.

Contributions to birdsong from the left and right sides of the intact syrinx

Roderick A. Suthers

School of Medicine and Department of Biology, Indiana University, Bloomington, Indiana 47405, USA

THE vocal organ, the syrinx, of some songbirds has been hypothesized to contain two sound sources that can be operated independently. The syrinx of songbirds (Oscines) is a bipartite structure^{1,2} whose two sides are potentially capable of acting either together or independently to produce sound^{3–9}. Sound production is lateralized in some species such that one side produces most of the song^{9–11}. I have now directly measured the acoustic output and motor dynamics of the left and right sides of the syrinx during song in catbirds and thrashers. In these birds, sound may be produced by either side of the syrinx alone, by both sides acting together, or by switching from side to side. When both sides of the syrinx contribute simultaneously to a note or syllable, both may generate the same sound or each side may produce a different sound. A given syllable type is generated by a similar motor pattern each time it is produced.

The rate of airflow and sound on each side of the syrinx was measured in three adult male grey catbirds (*Dumetella carolinensis*) and three adult male brown thrashers (*Toxostoma rufum*) (Mimidae) by implanting a microbead thermistor near the centre of the lumen of each primary bronchus a few millimetres caudal to the medial tympaniform membrane (Fig. 1a). The thermistors responded to respiratory airflow^{12,13} (Fig. 2) and to air oscillations up to about 3 kHz produced in the primary

Received 15 June; accepted 14 August 1990.

1. Pennycuik, C. J. *Ibis* **111**, 525–556 (1969).
2. Tucker, V. A. *J. exp. Biol.* **58**, 689–709 (1973).
3. Pennycuik, C. J. in *Avian Biology* (eds Farner, D. S. & King, J. R.) Vol. 5, 1–75 (Academic, London, 1975).
4. Greenewalt, C. H. *Trans. Am. phil. Soc.* **65**(4), 1–67 (1975).
5. Rayner, J. M. V. *J. exp. Biol.* **80**, 17–54 (1979).
6. Pennycuik, C. J. *Bird Flight Performance. A Practical Calculation Manual* (University Press, Oxford, 1989).
7. Tucker, V. A. *J. exp. Biol.* **48**, 67–87 (1968).

bronchus by the acoustic signal generated on the ipsilateral side of the syrinx (Fig. 1*b, c*).

Similar results were obtained from a fourth thrasher in which airflow through each side of the syrinx was measured by thermistors positioned in the caudal end of the trachea at the opening of each bronchus. This bird served as a control for the possible effect of bronchial thermistors on song. The large, variable song repertoires of these mimic thrushes¹⁴⁻¹⁶ and the limited sample recorded during experiments precluded a detailed comparison of syllable types before and after surgery.

About 10-23% of the syllables examined from each bird ($n=109-211$) were produced by only one side of the syrinx, either the left (Fig. 3*a*) or the right. The timing of phonation on each side was controlled by opening or closing the syringeal lumen. In 13-69% of the syllables studied, sound production switched from one side of the syrinx to the other, or to both,

during the course of the syllable as the two sides of the syrinx were alternately opened and closed. Some vocalizations contained only a single switch (Fig. 3*b*), but in others the bird repeatedly changed phonation back and forth between the left and right sides (Fig. 4*a*). Often the switch from one side to another occurred between successive notes of a multinote syllable but sometimes the bird changed sides within a single note.

Both sides of the syrinx generated sound simultaneously in 21-67% of the syllables studied. In many of these cases, both sides generated the same sound for the entire syllable. In other cases, the two sides generated sounds with independent frequency modulation and without a harmonic relationship to each other (Fig. 4*b*). These recordings provide the first direct experimental verification of the 'two voice' theory which Stein⁵ and Greenewalt⁶ proposed based on the acoustic properties of the song and the structure of the vocal organ. Sometimes these

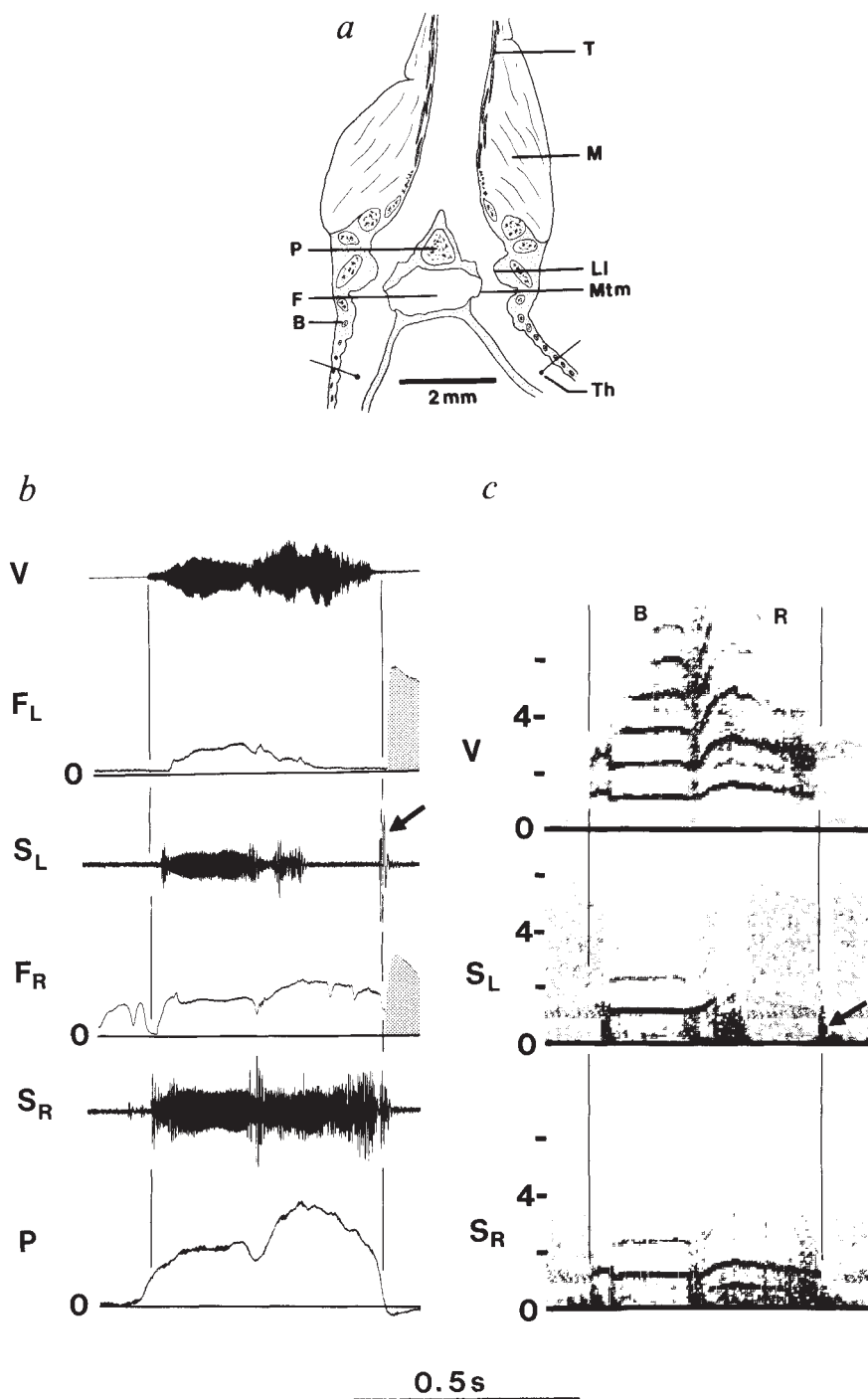


FIG. 1 *a* Frontal section through the syrinx of a brown thrasher, showing position of the thermistors used to record airflow and bronchial sound. T, tracheal ring; M, syringeal muscles (Innervated by the tracheosyringeal branch of CN XII); LI, lateral labium; Mtm, medial tympaniform membrane; Th, thermistor bead; P, pessulus; F, interbronchial foramen (continuous with the interclavicular air sac surrounding the syrinx); B, bronchial cartilage. *b*, Oscillographic display of a single note syllable from a catbird's song together with respiratory airflow and bronchial sound (obtained by high-pass filtering of the amplified thermistor output at 200 Hz) on each side of the syrinx. *c*, Spectrographic displays of the syllable and bronchial sound shown in *b*. Both sides are producing the same sound during the first half of the syllable, but the last half is produced only by the right side. Diagonal arrows indicate low-frequency acoustic transients often recorded in the bronchus during the opening or closure of the ipsilateral syrinx. These may be damped oscillations of the Mtm or LI associated with the initiation or termination of airflow before a phonatory configuration is achieved and are presumably filtered from the vocalization by the vocal tract. V, emitted vocalization recorded with microphone; F_L and F_R, rate of respiratory bronchial airflow on left and right side, respectively; S_L and S_R, sound in left and right bronchus, respectively; P, pressure in cranial thoracic air sac. R, L, and B, portions of notes or syllables generated on right, left or both sides of syrinx, respectively. Sound frequency is in kHz. Vertical lines align the beginning and end of vocalization; shading indicates inspiratory airflow.

FIG. 2 Five successive syllables from a catbird song. During full song, the contribution from each side of the syrinx, as indicated by airflow during phonation, varies from syllable to syllable. The song, recorded by a microphone in front of the cage, is shown spectrographically (top row) and oscillographically (fourth row). Singing began about two days after surgery and data were recorded for up to about one week, when the signal from thermistors began to deteriorate. Before an experiment, a 10 or 15 mg pellet of testosterone propionate was implanted subcutaneously to increase singing. Air sac pressure (P , cm H_2O), recorded by a piezoresistive pressure transducer connected to a cannula in the cranial thoracic air sac, closely approximates pressure in both primary bronchi²³. The positive subsyringeal pressure without airflow through the silent side of the syrinx confirms that it was closed during the syllable. Closure is presumably achieved by adduction of the lateral labium as, after unilateral neurotomy, birds were unable to regulate airflow through the operated side of the syrinx except by varying the air sac pressure. Regulation by the intact side remained normal. The valving mechanism controlling phonation is thus located in the syrinx, as opposed to the air sacs or bronchi⁹, and is controlled by the ipsilateral syringeal muscles. Leads from transducers were attached to the bird's back in a way that permitted it to move freely within its cage. Similar patterns of flow and pressure accompany the song of the brown thrasher. See legend of Fig. 1 for explanation of other symbols.

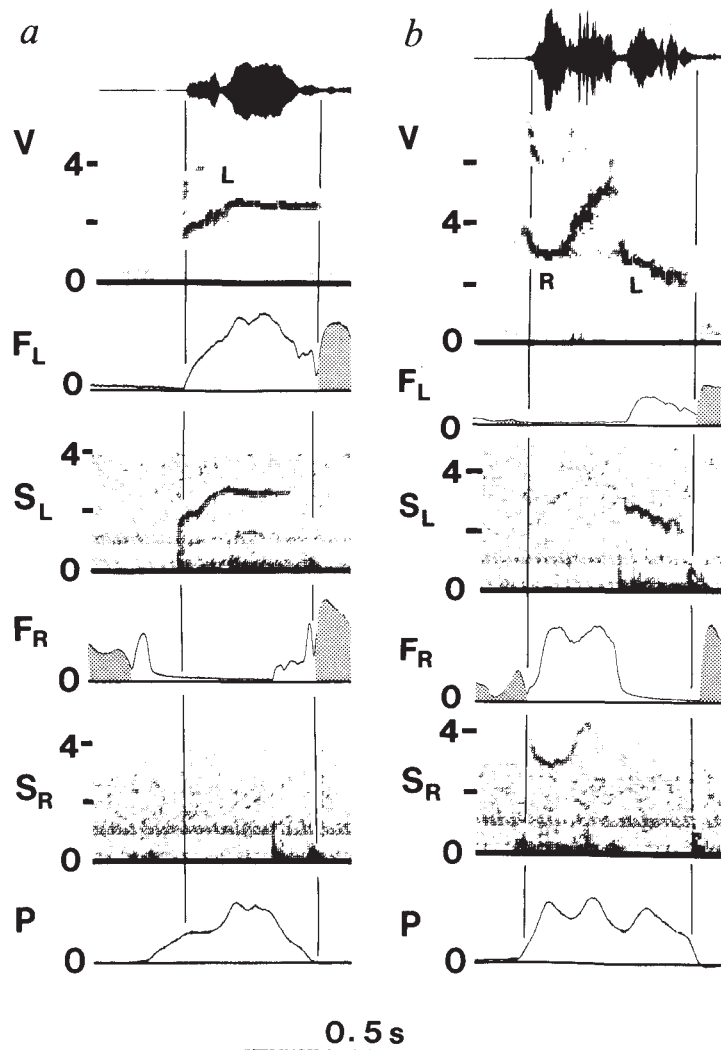
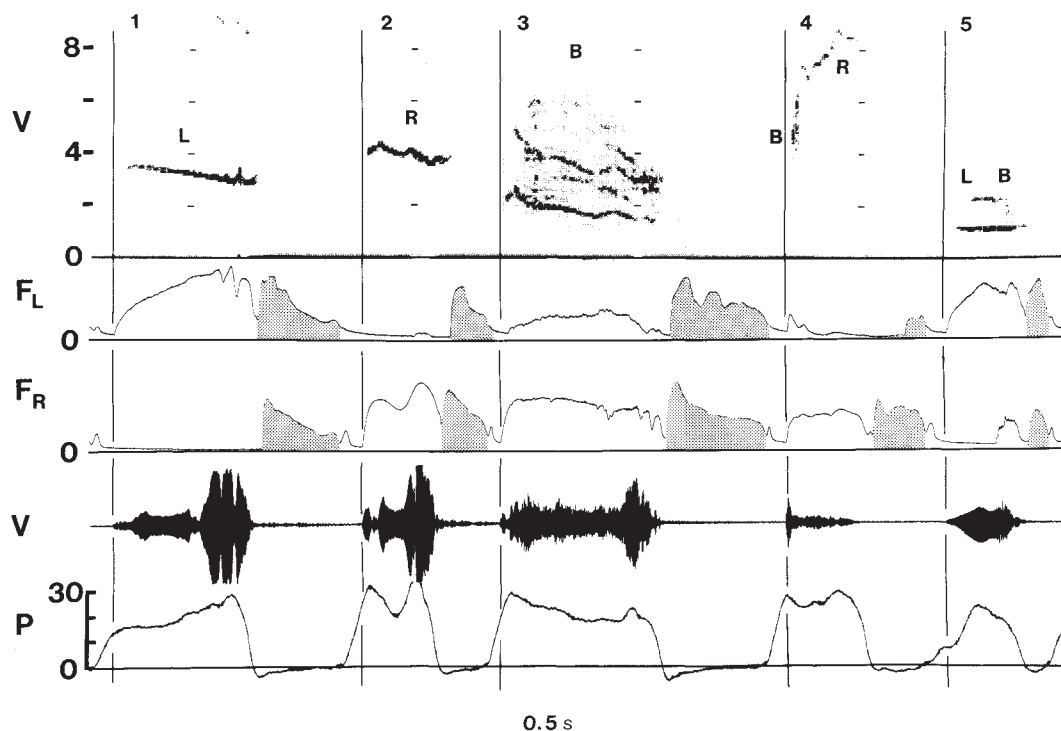


FIG. 3 *a*, Catbird single note syllable generated by the left syrinx while the right syrinx was closed. *b*, Catbird two-note syllable in which first note is generated by right syrinx, second note by left syrinx. The term 'note' refers to an elementary unit that appears as a single continuous marking on a sound spectrogram. Two or more notes may be combined to form a 'syllable' and single note syllables occur when individual notes are spaced like multinote syllables²⁴. The syllable spacing of catbird and thrasher song follows, with few exceptions, the respiratory rhythm, a single or multinote syllable being produced during each expiration. Faint continuous tone at about 1 kHz in recordings of bronchial sound is an electronic artefact. See legend of Fig. 1 for explanation of symbols.

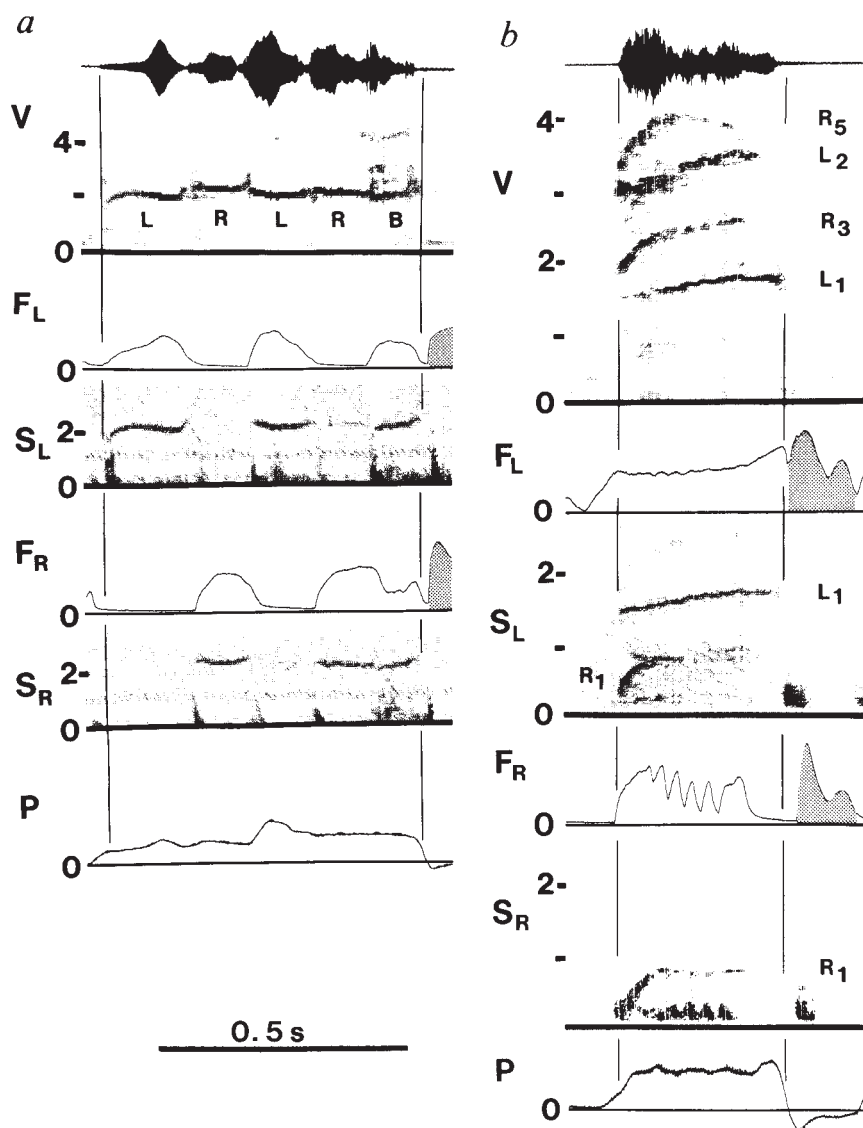


FIG. 4 *a*, Catbird syllable produced with alternating contributions from the left, right and finally both sides of the syrinx. The first notes are produced on different sides, the last note received sequential contributions from both sides. *b*, A catbird syllable in which the two sides of the syrinx are operating independently to produce two independent 'voices'. The emitted vocalization contains two prominent elements, each of which has harmonic components present, but which are not harmonically related. L_1 is a fundamental originating in the left syrinx together with its second harmonic L_2 . The right syrinx is generating a different fundamental, R_1 . R_1 and its even-numbered harmonics are suppressed in the emitted vocalization, presumably by vocal tract filtering. The attenuated version of the initial upward sweeping portion of R_1 that is also recorded in the left bronchus may be passively driven 'cross-talk' from the right side or actively generated on the left. See legend of Fig. 1 for explanation of other symbols.

'two voice' syllables were amplitude modulated at a periodicity equal to the difference frequency between the signals arising on each side, but no evidence of nonlinear interaction between separate harmonic series arising on each side of the syrinx, as occurs in the 'dee' syllable of chickadee (*Parus atricapillus*) song^{7,8}, was observed.

Neither catbirds nor thrashers as a group consistently favoured one side of the syrinx over the other. Among the four thrashers studied, the proportion of syllables produced by only the left versus only the right side was 21.3% versus 1.6%, 8.0% versus 12.3%, 5.6% versus 16.0%, and 0% versus 18.4%, respectively. In the case of the three catbirds these left-versus-right side proportions were 13.2% versus 0%, 11.8% versus 8.9%, and 3.1% versus 6.9%, respectively. The remaining majority of syllables were produced by both sides either simultaneously or sequentially. These birds lack a strong lateralization of song production like that reported in species such as canaries (*Serinus canaria*) and chaffinches (*Fringilla coelebs*)⁹⁻¹¹, in which one side of the syrinx generates most of the song. No obvious differences were apparent between sounds produced on different sides of the catbird or thrasher syrinx.

A given single or multinote syllable type was accompanied by a strikingly similar pattern of airflow and subsyringeal pressure that was as distinctive as the spectrographic representation

of the syllable it generated. The acoustic contribution from each side of the syrinx was likewise consistent. This was true for all the syllable types repeated by each bird (thrashers, 60; catbirds, 24), not only when the syllable was repeated immediately, as in a thrasher couplet, but also when either species repeated the syllable at a later time after many intervening notes. Each syllable type in the repertoire of these birds thus seems to be represented in the nervous system as a unique combination of stereotyped motor patterns for the two sides of its syrinx, as well as for the muscles of its respiratory system^{13,17} and vocal tract¹⁸. It remains to be determined whether different individuals use similar motor patterns to produce shared syllables.

It has been suggested that some birdsong may be generated aerodynamically by a constriction in the syringeal lumen that produces a series of stable vortices downstream, analogous to a hole-tone whistle¹⁹⁻²¹, rather than by vibrating membranes^{5,6,22}. Air oscillations recorded as sound in the bronchus, upstream from the syrinx, are most probably produced by a vibrating sound source, supporting the vibrating membrane hypothesis in the case of catbirds and thrashers. Syringeal structure varies greatly among taxa², however, and a whistle mechanism in other birds cannot be ruled out.

Recent evidence¹⁸ indicates that the vocal tract of some songbirds acts as a filter, altering the spectral composition of the

emitted vocalization by actively tracking the fundamental frequency and suppressing harmonics. In catbirds and thrashers, the occasional suppression of even-numbered harmonics (Fig. 4b) is presumably the result of vocal tract filtering. These birds also seem able to vary the filter properties of their vocal tract, but if they track a particular frequency component it is not always the fundamental since this is sometimes filtered from the vocalization. □

Received 12 April; accepted 19 July 1990.

1. Gaunt, A. S. in *Bird Respiration* Vol. 1 (ed. Seller, T. J.) 71–94 (CRC, Boca Raton, Florida, 1987).
2. King, A. S. in *Form and Function in Birds* Vol. 4 (eds King, A. S. & McLelland, J.) 105–192 (Academic, New York, 1989).
3. Potter, R. K., Kopp, C. A. & Green, H. C. *Visible Speech* (Van Nostrand-Reinhold, Princeton, New Jersey, 1947).
4. Borror, D. J. & Reese, C. R. *Ohio J. Sci.* **56**, 177–182 (1956).
5. Stein, R. C. *Auk* **85**, 229–243 (1968).
6. Greenewalt, C. H. *Bird Song: Acoustics and Physiology* (Smithsonian Inst., Washington DC, 1968).
7. Nowicki, S. & Capranica, R. R. *Science* **231**, 1297–1299 (1986).
8. Nowicki, S. & Capranica, R. R. *J. Neurosci.* **6**, 3595–3610 (1986).
9. Nottebohm, F. *J. exp. Zool.* **177**, 229–261 (1971).
10. Nottebohm, F. in *Lateralization in the Nervous System* (ed. Harnad, S., Doty, R. W., Goldstein, L., Jaynes, J. & Krauthamer, G.) 23–44 (Academic, New York, 1977).
11. Arnold, A. P. & Bottjer, S. W. in *Cerebral Lateralization in Nonhuman Species* (ed. Glick, S. D.) 11–39 (Academic, New York, 1985).
12. Suthers, R. A. & Hector, D. H. *J. comp. Physiol. A* **156**, 243–266 (1985).
13. Hartley, R. S. & Suthers, R. A. *J. comp. Physiol. A* **165**, 15–26 (1989).
14. Thompson, W. L. & Jane P. L. *Jack-pine Warbler* **47**, 115–125 (1969).
15. Boughey, M. J. & Thompson, N. S. *Z. Tierpsychol.* **56**, 47–58 (1981).
16. Kroodsma, D. E. & Parker, L. D. *Auk* **94**, 783–785 (1977).
17. Hartley, R. S. *Resp. Physiol.* (in the press).
18. Nowicki, S. *Nature* **325**, 53–55 (1987).
19. Gaunt, A. S., Gaunt, S. L. & Casey, R. M. *Auk* **99**, 474–494 (1982).
20. Casey, R. M. & Gaunt, A. S. *J. theor. Biol.* **116**, 45–64 (1985).
21. Gaunt, A. S. & Gaunt, S. L. in *Current Ornithology*. Vol. 2 (ed. Johnston, R. F.) 213–245 (Plenum, New York, 1985).
22. Miskimen, M. *Auk* **68**, 493–504 (1951).
23. Brackenbury, J. H. *Resp. Physiol.* **13**, 319–329 (1971).
24. Konishi, M. *A. Rev. Neurosci.* **8**, 125–170 (1985).

ACKNOWLEDGEMENTS. I thank R. Hartley, M. Konishi and F. Nottebohm for constructive suggestions, which have improved this manuscript, and R. Hartley for assisting in the experiments. This work was supported by the NSF.

Two components of long-term potentiation induced by different patterns of afferent activation

Lawrence M. Grover & Timothy J. Teyler

Neurobiology Department, Northeastern Ohio Universities College of Medicine, Rootstown, Ohio 44272, USA

LONG-TERM potentiation (LTP) of excitatory synaptic transmission^{1–4} could be a mechanism underlying memory^{3,5}. Induction of LTP requires Ca²⁺ influx into postsynaptic neurons⁶ through ion channels gated by NMDA (N-methyl-D-aspartate) receptors in hippocampus (area CA1^{7,8} and dentate gyrus⁹) and neocortex¹⁰. Here we report that a component of LTP not requiring the activation of NMDA receptors can be induced in area CA1. The component is dependent on tetanus frequency, requires increases in postsynaptic intracellular Ca²⁺ concentrations, and is suppressed by an antagonist of voltage-dependent Ca²⁺ channels.

Extracellular and intracellular recordings were obtained from CA1 pyramidal neurons, and excitatory post-synaptic potentials (e.p.s.p.s) were elicited by stimulation of Schaffer collateral/commissural fibers^{11,12}. Brief tetani of these afferents at 25 Hz (Fig. 1a) or 200 Hz (Fig. 1b) gave rise to LTP in all slices examined. The NMDA receptor antagonist D,L-2-amino-5-phosphonovaleate (D,L-APV, 50 μ M; also called D,L-2-amino-5-phosphopentanoate) completely blocked LTP induced by 25-Hz tetani (Fig. 1a), but only partially reduced LTP induced by 200 Hz tetani (Fig. 1b). This APV-resistant component of LTP had an unusual time course. In 50 μ M D,L-APV, 200-Hz tetani

produced an immediate post-tetanic potentiation (PTP) that typically decayed within 3–5 min (Figs 1b, e and 3b) followed by the gradual development of LTP over the next 10–15 min.

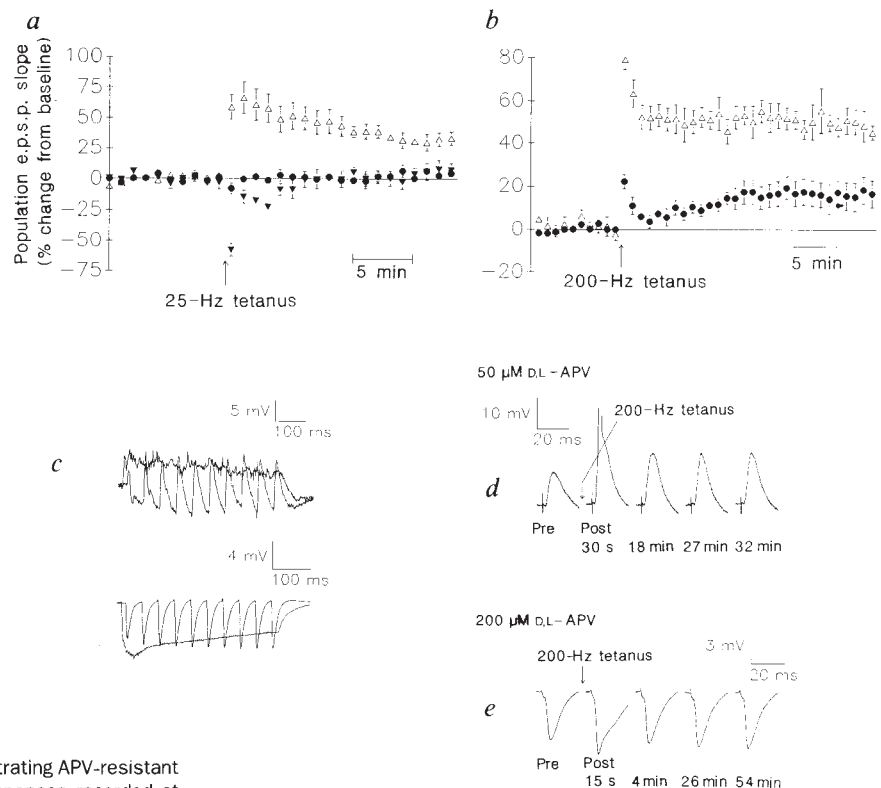
Tetanic stimulation releases large amounts of glutamate which might successfully compete with APV thus relieving the block of NMDA receptors. We used several approaches to test this possibility. First, intracellular recordings were obtained during 0.5-s, 200-Hz tetani. Figure 2a shows a depolarizing response recorded with APV present (trace i). The addition of DNQX (6,7-dinitroquinoxaline-2,3-dione) (Fig. 2a, trace ii) blocked synaptic potentials mediated by quisqualate (AMPA) or kainate receptors, leaving a small hyperpolarizing response that was blocked by the GABA_A receptor antagonist bicuculline (Fig. 2a, trace iii). Washout of APV (leaving DNQX and bicuculline) revealed a large NMDA receptor-mediated potential (Fig. 2a, trace iv), which was blocked by reperfusion of APV. We obtained similar results in five slices. These data indicate that 50 μ M D,L-APV is effective in blocking NMDA receptor-mediated synaptic responses during the 200-Hz tetani we used to induce LTP. We investigated this more fully in another set of slices by examining the dose-response relationship between APV concentration and depolarizations mediated by NMDA receptors during 200-Hz tetani. Results (Fig. 2b) indicated a near-complete block with 20 μ M D,L-APV, and no additional inhibition by up to 200 μ M D,L-APV. Second, in field potential experiments we increased the concentration of D,L-APV up to 200 μ M and still observed an APV-resistant component of LTP (Fig. 1e). Finally, to ensure that the increased number of stimuli used (400 stimuli at 200 Hz versus 25 stimuli at 25 Hz) was not responsible for these observations, we stimulated at 25 Hz for 16 s (400 stimuli) in 50 μ M D,L-APV. This resulted in depression lasting about 5 min, but no potentiation (Fig. 1a).

At most synapses, including CA1 (but not CA3¹³), LTP is thought to depend on activation of NMDA receptors^{7–10}. The finding of APV-resistant LTP in area CA1 is therefore surprising. Two factors may have allowed this observation. The first is the stimulation frequency during tetanization. In other studies, tetanus frequencies often range from 50 to 100 Hz, below the 200 Hz that we found effective. The second is the effect of distributing tetani across four short trains. Responses to 200-Hz tetani delivered in APV increase across repeated trains (Fig. 2c), which should facilitate the induction of LTP. It is also surprising that the same input pathway showed both NMDA receptor-dependent and -independent components of LTP. This is not the case in area CA3 where mossy fibre inputs show NMDA receptor-independent LTP, whereas commissural/associational inputs show NMDA receptor-dependent LTP¹³.

Could a route other than NMDA receptor-gated channels allow sufficient Ca²⁺ to enter CA1 neurons during tetanization to induce LTP? One possible route is through voltage-sensitive Ca²⁺ channels (VSCCs). Dihydropyridine sensitive Ca²⁺ channels are present in CA1 pyramidal neurons¹⁴. VSCCs¹⁵ and in particular dihydropyridine-sensitive Ca²⁺ channels¹⁶ can mediate synaptically activated Ca²⁺ entry into CA1 neurons. We therefore tested the ability of the dihydropyridine Ca²⁺ channel antagonist nifedipine to inhibit the NMDA receptor-independent component of LTP. Nifedipine (10 μ M) substantially reduced LTP induced in the presence of APV (Fig. 3). This suggests VSCCs participate in the induction of NMDA receptor-independent LTP. Sustained depolarization during tetanization (Fig. 1c) and across repeated tetani (Fig. 2c) should result in greater Ca²⁺ influx through VSCCs, which may explain the dependence on tetanus frequency shown by this component of LTP; 25-Hz tetani do not produce sustained depolarization, whereas 200-Hz tetani do (Fig. 1c).

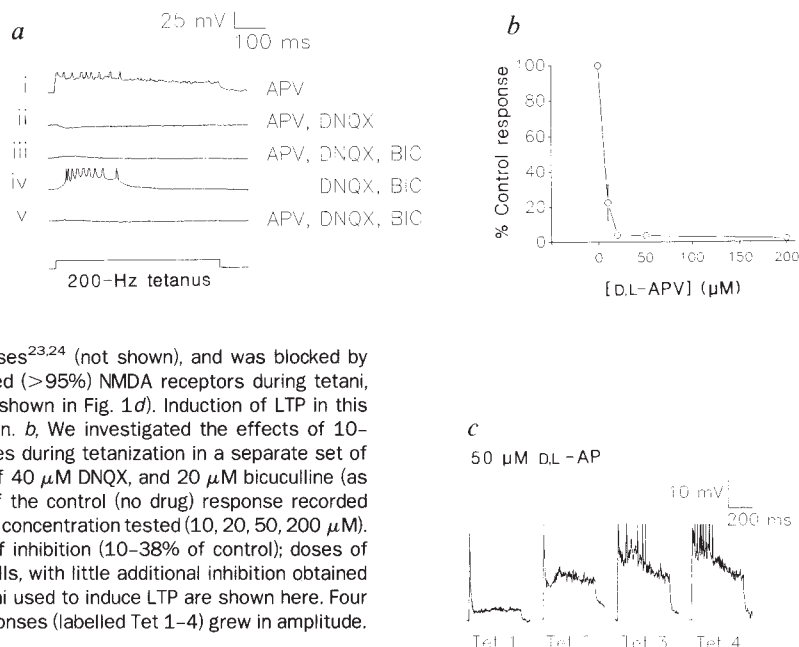
In some preparations, dihydropyridine-sensitive Ca²⁺ channels contribute to transmitter release^{17,18}. Because increased transmitter release may in part underlie expression of LTP¹⁹, nifedipine might have acted at a presynaptic site. To test if postsynaptic VSCCs participate in the induction of LTP by

FIG. 1 LTP in area CA1 of hippocampus after 25-Hz and 200-Hz tetani. *a*, *b*, Slope of e.p.s.p.s. is plotted as a function of time, one point per min. Each point is the mean (± 1 s.e.m.) response for all slices under the same conditions. *a*, 25-Hz tetani induced an immediate, large potentiation of e.p.s.p.s. (Δ). Potentiation decayed slowly over the first 10–15 min post-tetanus, then stabilized at a level significantly above baseline ($P < 0.001$, $n = 7$). Perfusion of 50 μ M D,L-APV completely blocked potentiation (P values > 0.05) induced by 25-Hz stimulation for 1 s ($\bullet$; 25 stimuli, $n = 6$) or 16 s (∇ ; 400 stimuli, $n = 4$). *b*, 200-Hz tetani induced an initial, large potentiation of e.p.s.p.s. (Δ), which showed some decay over the first 2–3 min post-tetanus (post-tetanic potentiation, PTP), and then remained stable (significantly greater than baseline, $P < 0.001$; $n = 6$). Perfusion of APV reduced, but did not eliminate LTP induced by 200-Hz tetani ($\bullet$; 400 stimuli total, $n = 8$). The remaining APV-insensitive component of LTP developed gradually over the 5–15-min period after tetanus. Responses recorded 25–30 min after tetanization significantly exceeded baseline ($P < 0.025$). Note the initial PTP that decayed within 2–3 min, before the development of the APV-insensitive component of LTP. *c*, Simultaneous intracellular (top) and extracellular (bottom) recordings from area CA1 during delivery of 25-Hz and 200-Hz tetani. The e.p.s.p.s. fuse and produce sustained depolarization during 200-Hz, but not 25-Hz, stimulation. *d*, Intracellular recording illustrating APV-resistant LTP. A pretetanus response is shown, as well as responses recorded at indicated times post-tetanus. *e*, Field potentials recorded from area CA1 in the presence of 200 μ M D,L-APV. PTP was present 15 s after tetanization, decayed within 5 min, and was followed by the gradual development of LTP. METHODS. Experiments were performed on *in vitro* 300- μ m-thick transverse hippocampal slices^{11,12}. Slices were continuously perfused with oxygenated, 35 °C medium containing (in mM): NaCl, 125.0; KCl, 3.35; NaH₂PO₄, 1.25; MgSO₄, 2.0; CaCl₂, 2.0; NaHCO₃, 25.0; glucose, 10.0. Some slices were perfused with medium containing 10–200 μ M D,L-APV. Extracellular field potentials were recorded in CA1 from stratum radiatum (s. rad) with glass micropipettes (2 M NaCl, 2–3 M Ω). Intracellular recordings were from CA1 pyramidal neurons impaled with glass micropipettes (4 M potassium acetate,



60–90 M Ω). Intracellular electrodes were connected to the headstage of a high input impedance electrometer with an active bridge circuit. Intracellular and extracellular potentials were amplified, digitized (10,000 Hz) and stored for later analysis. Test stimuli were delivered at a constant low rate (1–2 per minute) to s. rad. to activate Schaffer collateral/commissural inputs to CA1. Tetani were delivered as either a single 25-Hz, 1- or 16-s stimulus train (*a*) or four stimulus trains (200 Hz, 0.5 s) delivered at a rate of 1 train per 5 s (*b*). To determine if LTP was successfully induced under a particular set of conditions, matched sample *t*-tests were used to compare the mean response recorded during the baseline period with the mean response during the final 5 min of the recording.

FIG. 2 Intracellular recordings from CA1 pyramidal neurons during delivery of 200-Hz tetani. *a*, Depolarizing response recorded in the presence of 50 μ M D,L-APV (trace i) consists of a fused, sustained synaptic potential resulting from activation of kainate and/or quisqualate (K/Q) receptors. Action potentials (truncated) are superimposed on this response. DNQX (25 μ M) blocked the K/Q receptor-mediated synaptic depolarization, revealing a small hyperpolarizing response (trace ii). This hyperpolarization was blocked by subsequent addition of bicuculline (BIC, trace iii), indicating the involvement of GABA_A receptors. Removing APV from the perfusate (leaving DNQX and BIC) allowed a large depolarizing potential to develop during the tetanus (trace iv). This depolarization evoked a burst of action potentials, showed the voltage-dependence characteristic of NMDA receptor-mediated responses^{23,24} (not shown), and was blocked by reperfusion of APV (trace v). In this cell, 50 μ M D,L-APV blocked ($> 95\%$) NMDA receptors during tetani, but did not prevent development of a component of LTP (data shown in Fig. 1*d*). Induction of LTP in this neuron was therefore independent of NMDA receptor activation. *b*, We investigated the effects of 10–200 μ M D,L-APV on NMDA receptor mediated synaptic responses during tetanization in a separate set of slices. Intracellular responses were recorded in the presence of 40 μ M DNQX, and 20 μ M bicuculline (as illustrated in *a*, trace iv). Data are expressed as percentage of the control (no drug) response recorded from each cell. Shown here are the mean ± 1 s.e.m. ($n = 3$) for each concentration tested (10, 20, 50, 200 μ M). A 10 μ M concentration of D,L-APV produced varying degrees of inhibition (10–38% of control); doses of 20 μ M and higher resulted in 95% or greater inhibition in all cells, with little additional inhibition obtained by increasing concentration above 20 μ M. *c*, Responses to tetani used to induce LTP are shown here. Four tetani (200 Hz, 0.5 s) were presented every 5 s. Successive responses (labelled Tet 1–4) grew in amplitude. Same cell as shown in *a* and in Fig. 1*d*.



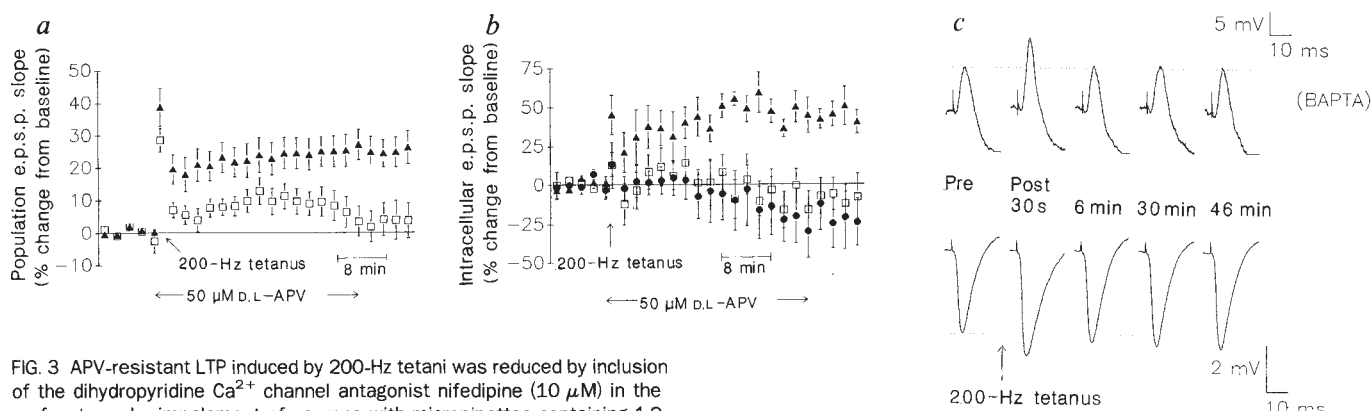


FIG. 3 APV-resistant LTP induced by 200-Hz tetani was reduced by inclusion of the dihydropyridine Ca^{2+} channel antagonist nifedipine ($10 \mu\text{M}$) in the perfusate, or by impalement of neurons with micropipettes containing 1,2-bis(2-aminophenoxy)ethane N,N,N',N' -tetraacetic acid (BAPTA). *a, b*, Each point is the mean (± 1 s.e.m.) response for all slices in the same condition, averaged over a 2-min period. *a*, Extracellular field e.p.s.p.s displayed APV-resistant LTP ($\blacktriangle$) when slices received 200-Hz tetani ($n=19$). Perfusion with $10 \mu\text{M}$ nifedipine significantly reduced ($P < 0.01$, $n=11$) APV-resistant LTP ($\square$). *b*, Intracellular e.p.s.p.s showed APV-resistant LTP ($\blacktriangle$, $n=7$), which was significantly reduced when $10 \mu\text{M}$ nifedipine ($\square$) was included in the perfusate ($P < 0.01$; $n=6$), or when BAPTA-filled micropipettes ($\bullet$) were used to impale cells ($P < 0.01$, $n=6$). *c*, Simultaneous intracellular and extracellular recordings illustrating block of APV-resistant LTP by Ca^{2+} chelator BAPTA. Intracellular e.p.s.p.s (top) recorded before and at indicated times after delivery of 200-Hz tetanus (in presence of $50 \mu\text{M}$ D,L-APV). The e.p.s.p.s showed a transient increase 30 s post-tetanus (PTP), which decayed over the next 2–3 min, and remained at the baseline level for the duration of the recording. The corresponding extracellular field e.p.s.p. (bottom) also showed a large PTP, but was followed by an APV-resistant LTP that developed

allowing Ca^{2+} influx into CA1 neurons, we impaled cells with microelectrodes filled with 1,2-bis(2-aminophenoxy)ethane N,N,N',N' -tetraacetic acid (BAPTA), a highly selective, rapid acting, Ca^{2+} chelator²⁰. These neurons failed to develop LTP (Fig. 3*b*). Increased intracellular Ca^{2+} concentration ($[\text{Ca}^{2+}]_i$) in postsynaptic neurons is therefore required for the APV-resistant component of LTP, as well as for the APV-sensitive component⁶. Dihydropyridine-sensitive Ca^{2+} channels may allow increases in $[\text{Ca}^{2+}]_i$ during tetanus. This agrees with a previous report showing enhancement of LTP by agonists of VSCCs²¹, but does not agree with a report showing no effect on LTP by antagonists of VSCCs²². In the latter report, however, LTP was induced by 100-Hz tetani, which may not be optimal for inducing this component of LTP.

The frequency dependency of the APV-resistant component of LTP suggests a possible function: a component of LTP induced only by an afferent input of higher frequency might allow the intensity of the potentiating input to be encoded. Although both components require increases in $[\text{Ca}^{2+}]_i$ for induction, the delayed onset of the APV-resistant component suggests some differences in the mechanisms of expression of LTP. $\square$

Received 6 June; accepted 3 August 1990.

- Bliss, T. V. P. & Gardener-Medwin, A. R. *J. Physiol.* **232**, 357–374 (1973).
- Bliss, T. V. P. & Lomo, T. *J. Physiol.* **232**, 331–356 (1973).
- Teyler, T. J. & DiScenna, P. A. *Rev. Neurosci.* **10**, 131–161 (1987).
- Kuba, K. & Kumamoto, E. *Prog. Neurobiol.* **34**, 197–269 (1990).
- Teyler, T. J. & DiScenna, P. *Brain Res. Rev.* **7**, 15–28 (1984).
- Lynch, G., Larson, J., Kelso, S., Barrionuevo, G. & Schottler, F. *Nature* **305**, 719–721 (1983).
- Collingridge, G. L., Kehl, S. J. & McLennan, H. *J. Physiol.* **334**, 33–46 (1983).
- Harris, E. W., Ganong, A. H. & Cotman, C. W. *Brain Res.* **323**, 132–137 (1984).
- Morris, R. G. M., Anderson, E., Lynch, G. S. & Baudry, M. *Nature* **319**, 774–776 (1986).
- Artola, A. & Singer, W. *Nature* **330**, 649–652 (1987).
- Grover, L. M. & Teyler, T. J. *Brain Res.* **506**, 53–61 (1990).
- Teyler, T. J. *Brain Res. Bull.* **5**, 391–403 (1980).
- Harris, E. W. & Cotman, C. W. *Neurosci. Lett.* **70**, 132–137 (1986).
- Cortes, R., Suparilai, M., Karobath, M. & Palacios, J. M. *Neurosci. Lett.* **42**, 249–254 (1983).
- Regehr, W. G., Connor, J. A. & Tank, D. W. *Nature* **341**, 533–537 (1989).
- Jones, R. S. G. & Heinemann, U. H. *Brain Res.* **418**, 257–266 (1987).
- Middlemass, D. N. & Spedding, M. *Nature* **314**, 94–96 (1985).
- Rane, S. G., Holtz, G. G. & Dunlap, K. *Plügers Arch.* **409**, 361–366 (1987).
- Dolphin, A. C., Errington, M. L. & Bliss, T. V. P. *Nature* **297**, 496–497 (1982).
- Tsien, R. Y. *Biochemistry* **19**, 2396–2404 (1980).

within 6 min of the tetanus. LTP persisted for the remainder of the recording. **METHODS.** Nifedipine stock solutions (100 mM) were prepared in DMSO and diluted to a final concentration of $10 \mu\text{M}$ (0.01% DMSO final concentration). Control slices and cells were perfused with normal medium or medium containing 0.01% DMSO. DMSO alone had no effect on baseline e.p.s.p.s or LTP. Nifedipine had no effect on baseline e.p.s.p.s. When BAPTA-filled micropipettes (20–100 mM BAPTA, tetra-potassium salt, in 4 M potassium acetate) were used to impale neurons, BAPTA was allowed to leak from the pipette until spike accommodation and the slow after-hyperpolarization were blocked^{25,26} (not shown). Recording of baseline synaptic responses began when these Ca^{2+} -activated K^{+} potentials were no longer present. To determine if nifedipine or BAPTA reduced the APV-insensitive component of LTP, independent samples *t*-tests were used to compare the magnitude of LTP obtained over the final 5 min of the recording session.

- Mulkeen, D., Anwyl, R. & Rowan, M. *J. Neurosci. Lett.* **80**, 351–355 (1987).
- Taube, J. S. & Schwartzkroin, P. A. *Brain Res.* **379**, 275–285 (1986).
- Mayer, M. L. & Westbrook, G. L. *J. Physiol.* **354**, 29–53 (1984).
- Nowak, L., Bregestovski, P., Ascher, P., Herbet, A. & Prochiantz, A. *Nature* **307**, 462–465 (1984).
- Madison, D. V. & Nicoll, R. A. *J. Physiol.* **354**, 319–331 (1984).
- Lancaster, B. & Nicoll, R. A. *J. Physiol.* **389**, 187–203 (1987).

ACKNOWLEDGEMENTS. We thank N. Lambert and P. DiScenna for helpful comments and suggestions. This work was supported by the Office of Naval Research.

The *scid* mutation in mice causes a general defect in DNA repair

G. M. Fulop & R. A. Phillips*

Division of Immunology and Cancer Research, Research Institute, Hospital for Sick Children and Departments of Medical Genetics and Immunology, University of Toronto, Toronto, Canada M5G 1X8

MICE homozygous for the *scid* mutation on chromosome 16 have a severe combined immune deficiency^{1,2} as a result of their inability to correctly rearrange their immunoglobulin and T-cell receptor genes^{3,4}. In *scid* mice, when precursors for B and T lymphocytes reach the stage of development requiring expression of these surface receptors, a defective recombinase system aberrantly cuts and rejoins the receptor gene segments greatly reducing the efficiency of producing functional receptors. As a result, most *scid* mice have no detectable B or T lymphocytes. We have demonstrated that the *scid* defect is not specific to lymphocyte development. Myeloid cells and fibroblasts from *scid* mice show a marked increase in sensitivity to ionizing radiation, indicating that the *scid* mutation leads to an inability to repair DNA damage induced by ionizing radiation as well as interfering with rearrangement of the immunoglobulin and T-cell receptor genes.

In a series of experiments designed to investigate the function of non-lymphoid haematopoietic cells in *scid* mice, we exposed

* To whom correspondence should be addressed.

scid mice and coisogenic normal C.B-17 mice to sublethal doses of γ -radiation to study endogenous spleen colonies derived from stem cells in spleen. Sublethal irradiation kills large numbers of stem cells and other haematopoietic cells and induces the proliferation and differentiation of the surviving stem cells⁵. Following sublethal doses of irradiation, the stem cells surviving in spleen proliferate and differentiate, forming small colonies of 1 to 10 million cells by day 10 after irradiation. Enumeration of these endogenous spleen colony forming stem cells (CFU-S) provides an estimate of the frequency of stem cells and can be used as a measure of radiation sensitivity of this population of stem cells⁵. We had demonstrated previously a normal frequency of CFU-S in *scid* mice⁶. In normal mice, endogenous spleen colonies begin to appear at ~600 cGy (centigreys) of whole-body irradiation. Generally, at doses greater than 800–900 cGy, no endogenous colony forming stem cells survive⁷. Surprisingly, we found that sublethal irradiation of *scid* mice with doses as low as 300 cGy leads to endogenous spleen colonies. At a dose of 300 cGy, an average of 13 spleen colonies were observed in one experiment. A dose of 600 cGy to normal C.B-17 mice was required to produce this number of endogenous spleen colonies. At doses greater than 500 cGy, no endogenous spleen colonies were observed in *scid* mice. In contrast, endogenous spleen colonies are observed in normal mice up to doses of 700 cGy. These observations provided the first clue that *scid* mice may have an increased sensitivity to ionizing radiation.

To obtain an accurate measure of the radiation sensitivity of *scid* mice and the normal controls, we determined radiation survival curves for the granulocyte/macrophage colony-forming units (CFU-GM) in the bone marrow of normal and *scid* mice. Cell suspensions of bone marrow cells from these mice were exposed to varying doses of ionizing radiation. The cells were then plated in appropriate tissue culture conditions to allow colony formation to occur⁸. Colonies were counted on day seven and the survival determined as the per cent colony-forming cells compared to an unirradiated control. Figure 1 shows the survival curves of CFU-GM from *scid* and normal mice. It is readily apparent that the CFU-GM from *scid* mice are significantly more sensitive to ionizing radiation than are CFU-GM from normal mice. Survival curves are usually characterized by slope and the extrapolation number. The slope is expressed as D_0 , the dose of radiation required to reduce survival to 37% in the exponential portion of the survival curve; the extrapolation number is the surviving fraction obtained by extrapolating the survival curve to zero dose⁵. Table 1 summarizes the D_0 and extrapolation number of the survival curves obtained for CFU-GM from several *scid* and normal mice. These data clearly show that *scid* mice are roughly twice as sensitive to ionizing radiation as are the coisogenic normal C.B-17 mice.

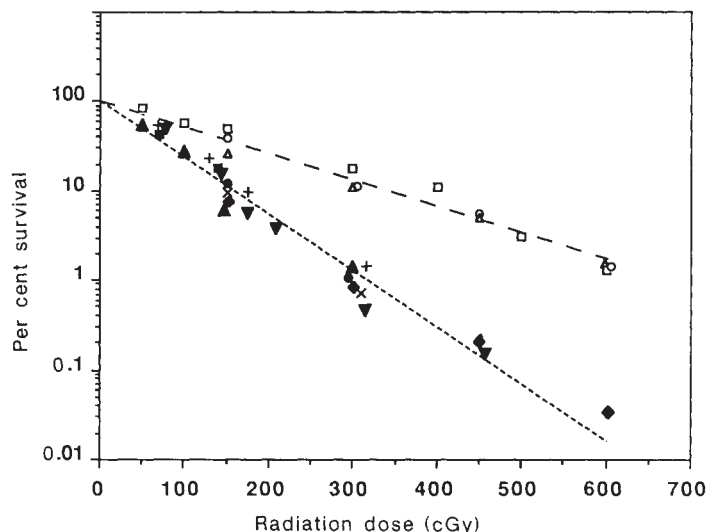


FIG. 1 Survival curves of CFU-GM from the bone marrow of *scid* and normal mice. The solid line is the average survival curve obtained from six independent determinations on cells from different *scid* mice; the six points shown (solid symbols) are the first six *scid* mice in Table 1. The dashed line is the average survival curve from three normal mice; the points are from the data summarized in Table 1; $\circ$ and Δ , first and third C.B-17, respectively, and $\square$, second BALB/cByJ.

METHODS. Survival curves were determined by harvesting bone marrow cells from mice and exposing the cell suspension to varying doses of γ -radiation from a ^{137}Cs irradiator²². The number of surviving CFU-GM were measured by plating the cells in methylcellulose with the appropriate growth factors as described previously^{8,23}. Colonies containing more than 50 cells were counted on day seven. The number of CFU-GM in the unirradiated suspension was defined as 100%; the number of CFU-GM in the irradiated suspensions was expressed as a percentage of the unirradiated control.

To confirm that increased radiation sensitivity is a direct effect intrinsic to the cells of the *scid* mouse, we compared the radiation sensitivity of *scid* mice, normal mice, and *scid* mice reconstituted with normal bone marrow as described previously⁹. Reconstituted animals have their myeloid and lymphoid systems totally derived from donor bone marrow¹⁰. The data in Table 1 show that CFU-GM in the bone marrow of reconstituted *scid* mice have a radiation sensitivity characteristic of normal mice. Thus, the increased sensitivity of such precursors in *scid* mice must be an intrinsic defect and not an anomaly in the microenvironment of the *scid* mouse.

Three experiments were done to confirm that the mutation leading to increased radiation sensitivity is identical to or closely linked to the *scid* mutation. First, we crossed *scid* mice to normal mice and then bred F1 hybrids to recover mice with the *scid* phenotype. Such *scid* mice from the F2 generation were identified by their lack of serum immunoglobulin. As shown in Table 1, *scid* mice obtained from this backcross experiment have a radiation sensitivity identical to that of *scid* mice obtained from the original inbred strain of *scid* mice. Second, *scid* mice were bred with M54 transgenic mice which are C57BL/6 mice carrying a rearranged immunoglobulin heavy chain gene¹¹. Transgenic *scid* mice were recovered by crossing the F1 mice with *scid* mice and screening the offspring for the transgene and for serum immunoglobulin. These *scid* mice also show increased radiation sensitivity (Table 1).

Third, *scid* mice were bred with normal mice and the resulting F1 mice crossed with *scid* mice to produce both normal and *scid* offspring. The radiation sensitivity of all of the mice was determined by measuring the survival of CFU-GM to 200 cGy of γ -radiation. These experiments are summarized in Table 2. It is clear from these results that the increased radiation sensitivity is recessive as both the F1 and immunoglobulin positive backcross mice show normal radiation sensitivity. As expected, the immunoglobulin negative *scid* progeny have increased sensitivity to radiation. As all of the *scid* mice in all three experiments have the same radiation sensitivity, we conclude that the increased sensitivity is closely linked to the *scid* mutation and is not the result of an independent mutation at a distant locus in the mouse genome. Also, the effect of the *scid* mutation on radiation sensitivity is recessive as is its effect on gene rearrangement.

Until our observations on the increased radiation sensitivity of CFU-GM, the only phenotype in *scid* mice was the absence of lymphocytes. As other genetic defects associated with increased radiation sensitivity involve defects in the enzymatic repair of DNA damage, the *scid* mutation may also involve a similar repair defect. To determine whether or not the defect in DNA repair leading to increased radiation sensitivity was gen-

eral or specific to haematopoietic cells, we examined the radiation response of kidney fibroblasts using the sensitive micronucleus assay^{12,13}. Micronuclei result from unrepaired chromosome breaks. When cells with such chromosome breaks go through mitosis, the small fragments of chromosome are left behind in the cytoplasm, forming small micronuclei. Enumeration of micronuclei is a good measure of radiation sensitivity. Figure 2 shows the number of micronuclei per 100 cells in fibroblasts from normal *scid*, backcross *scid*, and C.B-17 mice following exposure to different doses of ionizing radiation. As with the colony assay, it is clear that fibroblasts from *scid* mice are roughly twice as sensitive to ionizing radiation as are fibroblasts from normal mice. More importantly, this result confirms that the defect in DNA repair is general, affecting all cell types, and is not restricted to haematopoietic cells.

None of the human immune deficiencies is strictly analogous to the disease observed in *scid* mice. Ataxia telangiectasia (AT) is a rare autosomal recessive disorder characterized by several diverse clinical anomalies including enhanced radiation sensitiv-

ity^{14,15} and varying levels of immune deficiency¹⁶. Although the precise nature of the radiation repair defect in AT is unknown, it seems to be involved in some aspect of regulation of DNA synthesis. Normal cells, following exposure to ionizing radiation, shut down DNA synthesis to allow repair of radiation damage to occur in the absence of normal replication. In patients with AT there is an inability to shut down normal DNA synthesis¹⁴. Our preliminary data indicate that *scid* mice have the ability to shut down DNA synthesis in response to radiation damage. Thus, the *scid* mutation probably affects a different stage of repair from the AT mutation.

Our results clearly indicate that the *scid* mutation is pleiotropic, having effects on multiple organ systems. In the lymphoid system, the mutation results in a high frequency of aberrant rearrangements in the immunoglobulin and T-cell receptor genes. In both lineages, it has been suggested that the *scid* defect acts late during the recombination process^{3,17,18}. Pro-B cells in *scid* mice often correctly initiate cuts at the appropriate sites in the immunoglobulin gene and ligation occurs, but the cells splice together the cut ends incorrectly during gene rearrangement^{3,17}. The end result is a non-functional gene rearrangement. In other tissues, the *scid* defect manifests itself by a reduced ability to repair damage inflicted by ionizing radiation. If the same defective gene product affects gene rearrangement and repair of radiation damage, as suggested by our genetic analysis, one might suggest that the initial steps in radiation repair occur but that the mice are unable to ligate the ends created during the repair process correctly, thus leading to increased numbers of chromosome breaks and to increased lethality.

The results of these observations have several implications for the phenotype of the *scid* mouse. First, this mouse may show an increased susceptibility to carcinogens that require induction of mutations. It will be important to determine the spectrum of agents to which *scid* cells are unusually sensitive. It should be noted that genetic defects affecting repair of DNA damage are usually highly specific; thus patients with xeroderma pigmentosum are unable to repair ultraviolet damage but have normal sensitivities to X rays and most chemicals. Patients with AT cannot repair γ -radiation damage but are normal in their responses to ultraviolet and to many toxic chemicals^{14,19}. Patients with Fanconi's anaemia are highly sensitive to cross-linking agents such as mitomycin C but have normal sensitivity

TABLE 1 Summary of survival curves for *scid*, back-cross *scid*, reconstituted *scid* and normal mice

Mouse strain	D_0 (cGy)	n value	Mean values $\pm$ s.e.m.	
			D_0	n value
<i>scid</i>				
C.B-17 <i>scid</i>	77	0.67	65 ± 2.9	1.07 ± 0.08
	70	0.86		
	64	1.07		
	57	1.32		
	68	1.25		
	67	1.03		
	64	1.12		
	50	1.22		
BC <i>scid</i>	64	1.10	66 ± 3.57	1.18 ± 0.09
	68	0.98		
	57	1.40		
	74	1.22		
M54 <i>scid</i>	59	0.90	57	0.86
	56	0.81		
Reconstituted <i>scid</i> mice	136	0.97	147	0.98
	159	0.98		
Controls				
C.B-17	144	1.02	154 ± 8.80	1.08 ± 0.08
	184	1.07		
	150	0.88		
	132	1.35		
	161	1.07		
BALB/cByJ	161	0.99	155 ± 7.56	1.12 ± 0.07
	140	1.25		
	164	1.12		
C3H/HeJ	179	0.91	179	0.91

Each entry indicates results of a survival curve measurement made on cells from a single mouse. To determine the survival curves for CFU-GM, bone marrow from the indicated mice were exposed to varying doses of γ -radiation from a ¹³⁷Cs irradiator. The cells were cultured as described in the text to measure the number of surviving CFU-GM. D_0 is the radiation dose required to reduce the surviving fraction of cells to 0.37 in the exponential portion of the survival curve; n , the extrapolation number, is the surviving fraction obtained when the curve is extrapolated to zero dose. The original C.B-17 *scid* strain has been maintained and bred as a defined flora colony at the Ontario Cancer Institute since 1982. The BC or backcross strain was established in 1987 by backcrossing the original *scid* mice to coisogenic normal C.B-17 mice. The resulting F1 mice were mated and *scid* progeny identified by the absence of serum immunoglobulin. Brother and sister *scid* mice were then bred to establish a new *scid* line designated BC-strain. M54 *scid* mice were obtained by first crossing C.B-17 *scid* mice with M54 transgenic mice carrying a rearranged immunoglobulin heavy chain gene¹¹; transgenic *scid* mice were recovered by crossing the F1 with C.B-17 *scid* mice. The table shows the survival parameters for serum immunoglobulin negative, that is, *scid*, transgenic mice. Reconstituted *scid* mice received 400 cGy followed by 2×10^6 BALB/cByJ bone marrow intravenously; survival curves of CFU-GM were determined six weeks later.

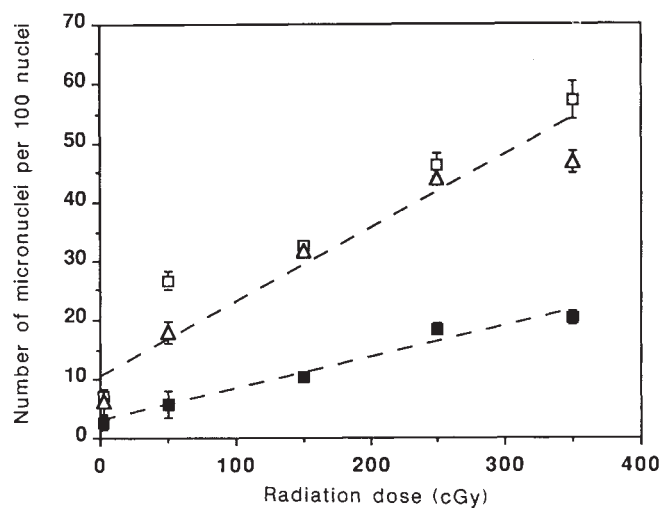
TABLE 2 Radiation sensitivity of *scid*, C.B-17, *scid* $\times$ normal F1, and backcross (*scid* $\times$ F1) mice

Mouse	Survival at 200 cGy (%)
<i>scid</i>	5.6
C.B-17	29
F1	52
	51
BC (Ig ⁺)	52
	59
	36
	50
	43
BC (Ig ⁻)	9.6
	4.6
	4.7
	6.0
	7.5

Normal bone marrow cells were prepared from the mice as indicated, exposed to 200 cGy of radiation, and surviving CFU-GM measured as described in the legend to Figure 1. The values shown for *scid* and for C.B-17 are the average values obtained for the experiments summarized in Table 1; other values indicate measurements made on separate individual mice. F1 refers to a cross between *scid* and normal mice and BC refers to offspring of *scid* $\times$ F1. (Ig⁺) denotes the immunoglobulin positive, that is, normal, offspring, and (Ig⁻) indicates the immunoglobulin negative, that is, *scid*, offspring.

FIG. 2 Production of micronuclei in fibroblasts from *scid* and normal mice following exposure to ionizing radiation. Four days after irradiation, the number of micronuclei per 100 nuclei were counted. The Figure shows the values obtained for fibroblasts from an inbred C.B-17 *scid* mouse (□), a backcross *scid* mouse (△) and a normal C.B-17 mouse (■). The points indicate the mean of the determinations made on the two slides for each mouse. The bars indicate the range of values measured; for points without bars, the range is smaller than the size of symbol.

METHODS. To establish fibroblast lines, kidneys from *scid* and normal mice were excised, minced and digested in 0.15% collagenase for 2 h. The mixture was then passed through sterile gauze and washed three times with Iscove's medium (IMDM)²³ plus 10% calf bovine serum (Hyclone). The cells were then seeded into a 75 cm² tissue culture flask in 5–10 mls of medium and incubated at 37 °C. After 24–48 h, all medium was removed and fresh medium added to the flask. Five days after the initial establishment of the culture, the adherent layer was trypsinized, washed and split for further growth. Once cells became confluent, cultures were split every 3–4 days. Determination of micronucleus frequency was carried out as previously described^{1,3}. In brief, flasks of exponentially growing lines from *scid* or C.B-17 mice were irradiated as described in the legend for Fig. 1 and incubated for 24 h at 37 °C in IMDM plus 10% calf bovine serum. The flasks were then trypsinized, collected and plated onto sterile slides in dilutions that would allow proliferation for the required time period. Slides were prepared for examination at the appropriate time intervals by removing the culture medium and exposing the slides to 75 mM KCl for 10 min, and then fixing in methanol:acetic acid (3:1) for 10 min. Slides were stained with Giemsa and the frequency of micronuclei for each cell line was estimated by counting micronuclei in 1,000–1,740 cells on two independent slides.



to X-ray and ultraviolet light. Until the response of *scid* cells to various DNA-damaging agents is known it will not be possible to accurately compare the *scid* mutation to the other known DNA repair defects.

Second, the pleiotropic effects of the *scid* mutation affect strategies that can be used to isolate the *scid* gene. It is clear from the data in this paper that all cells probably express the gene product of the *scid* locus. Its expression is unlikely to be limited to lymphoid tissues because all cells will need the *scid* gene product to repair DNA damage induced by various agents. Thus, it is not surprising that the recently identified, lymphocyte specific gene, *RAG-1*, involved in recombination of immunoglobulin genes²⁰ is located on a different chromosome than the *scid* locus¹.

Finally, it is interesting that the mechanism of gene rearrangement in lymphoid tissue has borrowed one of the enzymes involved in DNA repair. For this reason, it may be interesting to examine lymphocyte differentiation in various human diseases known to have defects in DNA repair. Although none of the

known DNA repair syndromes have immune deficiencies as marked as that observed in *scid* mice, many exhibit some immune defects. The identification of multiple complementation groups in many of the DNA repair syndromes indicates that many proteins are involved in the repair of DNA damage. It is likely that rearrangement of the immunoglobulin and T cell receptor genes also involves many proteins. On the basis of the results reported here, we should expect that some of the proteins will be used for both repair and rearrangement, as both mechanisms are similar in their manipulations of DNA. In addition to the processes involved in the rearrangement of germ-line genes to create functional lymphocyte receptors, B cells also rearrange the constant regions during the switch from IgM production to the production of IgG, IgA or IgE²¹. Proteins required for class switch may also be used by the DNA repair mechanism. Thus, abnormalities in the levels of different classes of immunoglobulin in patients with AT¹⁶ may reflect abnormalities in a common protein used both for DNA repair and for class switch during lymphocyte differentiation. □

Received 2 July; accepted 19 July 1990.

1. Bosma, G. C. et al. *Immunogenetics* **29**, 54–57 (1989).
2. Bosma, G. C., Custer, R. P. & Bosma, M. J. *Nature* **301**, 527–530 (1983).
3. Malynn, B. A. et al. *Cell* **54**, 453–460 (1988).
4. Schuler, W. et al. *Cell* **46**, 963–972 (1986).
5. Till, J. E. & McCulloch, E. A. *Radiat. Res.* **14**, 213–222 (1961).
6. Dorshkind, K. et al. *J. Immunol.* **132**, 1804–1808 (1984).
7. Till, J. E. & McCulloch, E. A. *Ann. N.Y. Acad. Sci.* **114**, 115–125 (1964).
8. Magli, M. C., Dick, J. E., Huszar, D., Bernstein, A. & Phillips, R. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 789–793 (1987).
9. Fulop, G. M. & Phillips, R. A. *J. Immunol.* **136**, 4438–4443 (1986).
10. Fulop, G. M., Wu, D. D. & Phillips, R. A. in *The Scid Mouse. Characterization and Potential Uses* (eds Bosma, M. J., Phillips, R. A. & Schuler, W.) 171–180 (Springer, Heidelberg, 1989).
11. Grosschedl, R., Weaver, D., Baltimore, D. & Constantini, F. *Cell* **38**, 647–658 (1984).
12. Countryman, P. I. & Heddle, J. A. *Mutat. Res.* **41**, 1343–1347 (1976).
13. Goddard, A. D., Heddle, J. A., Gallie, B. L. & Phillips, R. A. *Mutat. Res.* **152**, 31–38 (1985).
14. Paterson, M. C., Bech-Hansen, N. T., Smith, P. J. & Mulvihill, J. J. in *Radiation Carcinogenesis: Epidemiology and Biological Significance* (eds Boice, J. D., Jr & Fraumeni, J. F., Jr) 319–336 (Raven, New York, 1984).

15. McKinnon, P. J. *Mum. Genet.* **75**, 197–208 (1987).
16. Roifman, C. M. & Gelfand, E. W. in *Ataxia-Telangiectasia. Genetics, Neuropathology, and Immunology of a Degenerative Disease of Childhood* (eds Gatti, R. A. & Swift, M.) 273–285 (Alan R. Liss, New York, 1985).
17. Kim, M.-G., Schuler, W., Bosma, M. J. & Marcu, K. B. *J. Immunol.* **141**, 1341–1347 (1988).
18. Tutt, M. et al. *J. Immunol.* **138**, 2338–2344 (1987).
19. Arlett, C. F., Harcourt, S. A., Teo, I. A. & Broughton, B. C. in *Ataxia Telangiectasia—A Cellular and Molecular Link Between Cancer, Neuropathology and Immune Deficiency* (eds Bridges, B. A. & Harnden, D. G.) 347–353 (Wiley, London, 1982).
20. Schatz, D. G., Oettinger, M. A. & Baltimore, D. *Cell* **59**, 1035–1048 (1989).
21. Blackwell, T. K. & Alt, F. W. in *Molecular Immunology* (eds Hames, B. D. & Glover, D. M.) 1–60 (IRL, Oxford, 1988).
22. Cunningham, J. R., Bruce, W. R. & Webb, H. P. *Phys. med. Biol.* **10**, 318–332 (1965).
23. Iscove, N. N., Guilbert, L. J. & Weyman, C. *Expl. Cell. Res.* **126**, 121–126 (1980).

ACKNOWLEDGEMENTS. We thank S. Chance and C. Spencer for technical assistance. This research was supported by the Medical Research Council and the National Cancer Institute of Canada.

Small rearrangements in structures of Fv and Fab fragments of antibody D1.3 on antigen binding

T. N. Bhat, G. A. Bentley, T. O. Fischmann, G. Boulot & R. J. Poljak

Unité d'Immunologie Structurale (URA 359 CNRS), Département d'Immunologie, Institut Pasteur, 75724 Paris Cedex 15, France

THE potential use of monoclonal antibodies in immunological, chemical and clinical applications has stimulated the protein engineering and expression of Fv fragments¹⁻⁸, which are heterodimers consisting of the light and heavy chain variable domains (VL and VH) of antibodies. Although Fv fragments exhibit antigen binding specificity and association constants similar to their parent antibodies or Fab moieties, similarity in their interactions with antigen at the level of three-dimensional structure has not been investigated. We have determined the high-resolution crystal structure of the genetically engineered FvD1.3 fragment¹ of the anti-hen egg-white lysozyme (HEL) monoclonal antibody D1.3 (ref. 9), and of its complex with HEL. On comparison with the crystallographically refined FabD1.3-HEL complex, we find that FvD1.3 and FabD1.3 make, with minor exceptions, very similar contacts with the antigen. Furthermore, a small but systematic rearrangement of the domains of FvD1.3 occurs on binding HEL, bringing the contacting residues closer to the antigen by a

mean value of about 0.7 Å for VH (aligning on VL) or of 0.5 Å for VL (aligning on VH). This is indicative of an induced fit rather than a 'lock and key' fit to the antigen.

Preparation and crystallization of FvD1.3, FvD1.3-HEL and FabD1.3-HEL have been reported^{10,11}. X-ray diffraction data for FvD1.3, FvD1.3-HEL and FabD1.3-HEL were measured to 1.9 Å, 2.4 Å and 2.5 Å resolution, respectively. The structure of FabD1.3-HEL was refined starting from the 2.8 Å resolution model¹². The structures of FvD1.3 and FvD1.3-HEL were subsequently determined by molecular replacement using the entire Fv part from the FabD1.3 complex. After crystallographic refinement using the program X-PLOR¹³, the three structures gave *R* factors of 19% with root-mean-square (r.m.s.) deviations from ideal bond lengths of 0.017 Å or smaller. Details of the structure determinations will be published elsewhere.

The structure of the bacterial FvD1.3 in its complex with HEL closely superposes with the eukaryotic VH and VL domains in the FabD1.3-HEL complex, giving an r.m.s. difference in *C* α positions of 0.39 Å. In particular, a very close correspondence is seen in the position and orientation of the antigen-contacting residues (Fig. 1). Table 1 shows that FvD1.3 makes almost identical interatomic contacts with HEL as FabD1.3; the r.m.s. difference in these contact distances is 0.4 Å. Nonetheless, as shown in Table 1, some differences do exist. Similarly, the structure of the antigen is largely conserved in the two complexes. Although intermolecular contacts in the FvD1.3-HEL crystals affect the conformation of HEL residues 100-104, these residues are not involved in antigen-antibody contacts. The *C* α positions of the HEL moieties of the two complexes superpose (excluding residues 100-104 in the deformed loop) with an r.m.s. deviation of 0.39 Å.

TABLE 1 Interatomic distances between the antigen and the antibody

FvD1.3	VL	HEL		Å	Fab-Fv	FvD1.3	VH	HEL		Å	Fab-Fv	FvD1.3	VH	HEL		Å	Fab-Fv
Y32	C γ	Q121	N ϵ 2	3.5	-0.2	G31	C α	G117	C α	4.0	0.5	D100	O δ 1	S24	N	3.2	-0.4
Y32	C δ 1	Q121	N ϵ 2	3.6	-0.1	G31	C	G117	C α	3.6	0.3	D100	O δ 1	S24	C α	4.4	-0.8
Y32	C ϵ 1	Q121	N ϵ 2	3.5	0.1	G31	O	G117	C α	3.3	0.4	D100	O δ 1	S24	C β	4.4	-1.2
Y32	C δ 2	Q121	N ϵ 2	3.4	-0.1	Y32	C ζ	K116	N ζ	3.7	1.1	D100	O δ 1	S24	O γ	4.6	-1.6
Y32	C ϵ 2	Q121	N ϵ 2	3.3	0.1	Y32	O η	K116	N ζ	3.2	0.7	D100	O δ 2	S24	C β	3.4	0.4
Y32	C ζ	Q121	C γ	4.2	-0.2	W52	C ϵ 2	D119	C β	4.1	0.2	D100	O δ 2	S24	O γ	3.0	-0.4
Y32	C ζ	Q121	N ϵ 2	3.4	0.1	W52	C ϵ 3	G117	O	3.7	0.9	D100	O δ 2	N27	N δ 2	3.2	-0.2
Y50	C δ 2	N19	C β	3.9	0.5	W52	C ϵ 3	T118	C	3.8	0.1	Y101	C ϵ 1	V120	C γ 2	4.2	-0.3
Y50	C δ 2	N19	C γ	3.9	0.2	W52	C ϵ 3	D119	N	3.6	0.2	Y101	C ϵ 1	Q121	C β	4.0	-0.1
Y50	C δ 2	N19	N δ 2	3.7	0.0	W52	C ϵ 3	D119	C α	4.0	0.0	Y101	C ϵ 2	Q121	C β	4.0	0.0
Y50	C ζ	D18	C γ	4.1	0.0	W52	C ζ 2	D119	C β	3.7	0.3	Y101	C ζ	Q121	C β	3.4	0.1
Y50	C ζ	D18	O δ 1	3.7	0.2	W52	C ζ 2	D119	C γ	3.8	0.3	Y101	O η	D119	C α	3.5	-0.1
Y50	C ζ	D18	O δ 2	3.7	-0.1	W52	C ζ 3	T118	C	3.8	0.3	Y101	O η	D119	O δ 1	2.8	0.2
Y50	O η	D18	C γ	3.4	0.2	W52	C ζ 3	D119	N	3.5	0.4	Y101	O η	D119	C	3.5	-0.2
Y50	O η	D18	O δ 1	3.2	0.4	W52	C ζ 3	D119	C α	3.5	0.2	Y101	O η	V120	N	3.5	-0.2
Y50	O η	D18	O δ 2	2.8	0.1	W52	C ζ 3	D119	C β	3.7	0.3	Y101	O η	Q121	N	3.0	0.0
T53	O γ 1	N19	N δ 2	2.9	0.2	W52	C η 2	D119	C α	3.7	0.3	Y101	O η	Q121	C α	3.5	0.2
F91	O	Q121	C δ	3.7	0.1	W52	C η 2	D119	C β	3.5	0.3	Y101	O η	Q121	C β	2.9	0.5
F91	O	Q121	N ϵ 2	2.9	0.2	W52	C η 2	D119	C γ	3.8	0.1	R102	N η 1	G22	C	3.6	0.6
W92	C α	Q121	N ϵ 2	4.2	-0.4	W52	C η 2	D119	O δ 1	3.6	0.1	R102	N η 1	G22	O	2.5	0.7
W92	C ϵ 2	R125	C γ	4.0	0.3	G53	N	G117	O	2.9	0.0						
W92	C ϵ 3	Q121	C δ	3.5	0.0	G53	C α	G117	O	3.4	-0.2						
W92	C ϵ 3	Q121	O ϵ 1	3.5	0.1	G53	C	G117	O	4.2	-0.5						
W92	C ϵ 3	Q121	N ϵ 2	3.5	-0.5	D54	N	G117	O	3.8	-0.6						
W92	C ζ 2	R125	C γ	3.4	0.2	D54	C β	T118	C γ 2	3.7	0.6						
W92	C ζ 2	R125	C δ	4.2	-0.2	D54	C γ	T118	C γ 2	4.0	0.2						
W92	C ζ 3	Q121	C γ	4.0	-0.1	D54	O δ 2	T118	C γ 2	3.9	-0.3						
W92	C ζ 3	Q121	C δ	3.6	0.0	D100	C γ	S24	N	3.8	0.0						
W92	C ζ 3	Q121	N ϵ 2	3.8	-0.4	D100	C γ	S24	C β	4.2	-0.3						
W92	C ζ 3	I124	C δ 1	3.9	0.0	D100	C γ	S24	O γ	4.1	-0.9						
W92	C η 2	I124	C δ 1	3.9	0.0	D100	C γ	N27	N α 2	3.7	-0.2						
W92	C η 2	R125	C γ	3.7	-0.2	D100	O δ 1	G22	O	3.1	0.3						
S93	N	Q121	O ϵ 1	3.2	0.0	D100	O δ 1	Y23	C α	3.5	0.2						

Maximum contact distances of C-C: 4.1 Å, C-N: 3.8 Å, C-O: 3.7 Å, O-O: 3.3 Å, O-N: 3.4 Å, N-N: 3.4 Å (as in refs 23 and 24) occurring in the FvD1.3-HEL or the FabD1.3-HEL complexes are listed. The column labelled Fv contains those for FvD1.3-HEL. The column labelled Fab-Fv gives the values to be added to the FvD1.3-HEL contact distances to obtain those corresponding to FabD1.3-HEL. The mean and r.m.s. values of the differences in contact distances (column Fab-Fv) are 0.03 Å and 0.4 Å, respectively.

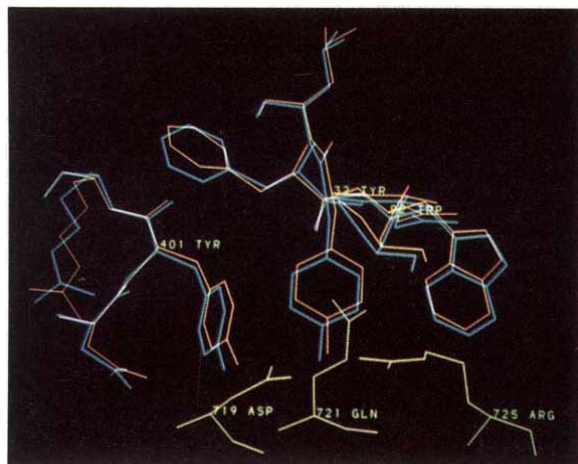


FIG. 1 The antigen-antibody interface around HEL Gln 121. Residues of FabD1.3 (orange) and FvD1.3 (blue) from their crystal complexes with HEL were superposed as described in Table 2. Gln 121 (labelled 721 in this figure) is surrounded by the antibody aromatic residues VL Tyr 32 and Trp 92 and VH Tyr 101 (labelled 401) and by the salt-bridged HEL residues Asp 119 (labelled 719) and Arg 125 (labelled 725). The side-chain atoms, N ϵ 2 and O ϵ 1, of Gln 121 make hydrogen bonds with O of Phe 91(VL) and N of Ser 93(VH) respectively, and in addition, the N ϵ 2 is favourably placed for a hydrogen bond interaction with the aromatic ring¹⁹ of Tyr 32(VL) (see Table 1). This figure illustrates the close correspondence of FabD1.3-HEL and FvD1.3-HEL structures at the combining site.

To study the effect of antigen binding on FvD1.3 and FabD1.3, comparisons were made between the free FvD1.3 and the antigen-bound FvD1.3 and FabD1.3 structures (Table 2 and Fig. 2). Superposition of the C α coordinates of the free FvD1.3 VL domain with the antigen-bound FvD1.3 VL domain gives an r.m.s. deviation of 0.37 Å, comparable to the estimated error¹⁴ in atomic coordinates (0.33 Å). A similar observation is made on superposing the VH domains, indicating that antigen binding does not introduce major conformational changes in the individual V domains. Indeed, this is borne out on comparing individual main-chain and side-chain conformations of both V domains in the three crystal structures. The effect of antigen binding on the relative arrangement of the V domains was also studied by superposing one of the domains and calculating the differences in position of the other. When superposing the VL domains of FvD1.3 and FvD1.3-HEL, for example, we find a small relative displacement of the VH domains (r.m.s. 0.99 Å, Table 2) which is larger than the corresponding movement observed between the FvD1.3-HEL and the FabD1.3-HEL structures (r.m.s. 0.52 Å). This systematic displacement moves VH closer to HEL such that the antigen-antibody contact distances (as listed in Table 1) for this domain decrease by a mean value of 0.72 Å. Alternatively, if the VH domains are superposed, then the VL domain contact distances with antigen decrease by a

mean value of 0.53 Å.

Although different relative arrangements of the variable domains of immunoglobulins have been reported¹⁵⁻¹⁷ (reviewed in ref. 18) crystallographic data associating such differences with antigen binding rather than antibody diversity were lacking. The change in the relative position of the variable domains reported here is based on a direct comparison of the structures of the free FvD1.3 and antigen-bound FvD1.3 and FabD1.3. It not only results in a closer approach of antibody contacting residues to the antigen, but in an increased separation of the variable domains from each other at their C termini (Fig. 2).

The study of these three crystal structures makes clear the importance of careful refinement with high resolution diffraction measurements to see the fine atomic details of antigen specificity. In the avian egg-white lysozymes from partridge, turkey, California quail, pheasant and Guinea hen, which are not bound⁹ by monoclonal antibody D1.3, HEL Gln 121 is replaced by His. In the complex, Gln 121 is surrounded by VL Tyr 32, Trp 92 and VH Tyr 101, and towards the solvent side by Arg 125 and Asp 119 of HEL, which are linked to each other by a salt bridge (Fig. 1). If Gln 121 were replaced by His, the imidazole side chain would make a number of disallowed, close van der Waal's contacts with the side chains mentioned above. Consequently, a His residue would not fit into the cavity occupied by the side

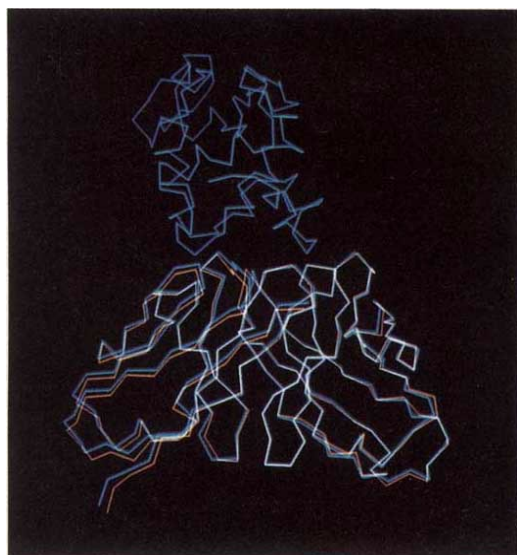


FIG. 2 Superposed C α backbones of FvD1.3 (orange) and the FvD1.3-HEL complex (blue). The C α atoms of the VH domains were superimposed as described in Table 2. The superposition shows the small but systematic displacement of the VL domain. Thus, the VL antigen-contacting residues Tyr 32, Tyr 50 and Ser 93 are displaced by 0.6 Å, 0.4 Å and 0.7 Å respectively towards the antigen, whereas at the antigen-distal part, displacements of 1.4 Å (Gly 16), 1.1 Å (Pro 59), 1.3 Å (Ser 77) are observed. Similarly, we find that if the superpositions are based on VL (not shown), then the VH antigen-contacting residues Tyr 32, Asp 54, Tyr 101 are displaced by 0.8 Å, 1.3 Å and 0.9 Å respectively, and at the antigen-distal parts Pro 141, Ser 184 and Thr 187 are displaced by 0.6 Å, 1.8 Å and 0.9 Å, respectively.

METHODS. Crystallization, symmetry and unit cell dimensions of FvD1.3 and FvD1.3-HEL have been reported¹⁰. X-ray intensities from FvD1.3-HEL crystals were collected with a Siemens-Xentronics area detector and a rotating anode generator, and processed using the program XDS²⁰; 23,130 observations to 2.4 Å resolution were merged and scaled giving 13,045 independent reflections (89% complete) with a merging agreement of 9%. Refinement of atomic coordinates was done using the programs XPLOR¹³ and FRODO^{21,22}. Twenty water molecules were included in the final model. The final *R* factor is 0.19, with an r.m.s. deviation of 0.014 Å from ideal bond lengths. Assessment¹⁴ of the intrinsic accuracy of the atomic coordinates gives a mean error of 0.33 Å. X-ray intensity measurements of FvD1.3 were as for FvD1.3-HEL; 151,683 observations to 1.9 Å resolution were merged and scaled to give 17,905 independent reflections (76% complete); merging agreement is 15.5%. Refinement of atomic coordinates carried out as for FvD1.3-HEL. Twenty-four water molecules were added to the model. Final *R* factor is 0.19, with an r.m.s. deviation of 0.017 Å from ideal bond lengths. Intrinsic accuracy¹⁴ of coordinates is estimated to be 0.33 Å.

TABLE 2 Comparison of the relative arrangements of the V_H, V_L domains in free FvD1.3, the FvD1.3-HEL and FabD1.3-HEL complexes

	FvD1.3	FvD1.3-HEL	FabD1.3-HEL
FvD1.3	—	0.37 (0.99)	0.37 (0.88)
FvD1.3-HEL	0.51 (0.73)	—	0.30 (0.52)
FabD1.3-HEL	0.41 (0.88)	0.40 (0.45)	—

The α coordinates of the V_L (or V_H) domains of the FvD1.3, the FvD1.3-HEL and FabD1.3-HEL complexes were superposed by least squares to minimize distances between homologous α positions. The α positions of V_L have been superposed and the resulting r.m.s. deviations in Å for the α of V_L (in parentheses for the α for V_H) are given above (to the right) the dashes along the diagonal. The FvD1.3-HEL and FabD1.3-HEL give the best agreement (0.30 Å). The α positions of V_H have been superposed and the resulting r.m.s. deviation for the α of V_H (in parentheses for the α of V_L) are given below (to the left) the dashes along the diagonal. The best agreement (0.40 Å) is again observed between FvD1.3-HEL and FabD1.3-HEL, in spite of a local change of conformation in residues 12–18 of the FvD1.3-HEL V_H domain arising from intermolecular contacts in the crystal.

chain of Gln 121. The very low affinity of antibody D1.3 for heterologous antigens in which position 121 is occupied by His is thus due to steric hindrance. In addition, the replacement His would make at most one hydrogen bond instead of the three hydrogen bonds made by the side chain of Gln 121 (Fig. 1, and Table 1). Well-ordered, buried or partially buried water molecules mediate contacts between antigen and antibody. Examples of these are a water molecule placed between V_L Tyr 32, V_L Tyr 50 and HEL Asp 18, and another placed between V_H Tyr 32 and HEL Lys 116. These water molecules are observed in the electron density maps of both the FabD1.3-HEL and the FvD1.3-HEL complexes.

In summary, the FvD1.3 expressed in bacteria reacts with its specific antigen, making similar interatomic contacts to those of the eukaryotic FabD1.3. The high resolution D1.3 structures presented here allow a precise description of the antigen-antibody interface. The discriminating specificity of antibody D1.3 is based on the closely matching complementarity to the antigenic determinant, which can only be defined by high-resolution structure determinations as presented here. The comparison of free with antigen-bound FvD1.3 clearly indicates a relative displacement of the variable domains on binding HEL, giving rise to an induced fit to the antigen. This feature of antigen-antibody reactions has implications for predicting their interactions on the basis of the structures of the free reactants. □

Secretion of immunoglobulin M assembly intermediates in the presence of reducing agents

Cristina M. Alberini*, Paola Bet†, Cesar Milstein‡ & Roberto Sitia†§

* Istituto di Chimica, Facoltà di Medicina, Università degli Studi di Brescia, Brescia, Italy

† Istituto Nazionale per la Ricerca sul Cancro, Genova, Italy

‡ MRC Laboratory of Molecular Biology, Cambridge, UK

§ Istituto Scientifico San Raffaele, Via Olgettina 60, I-20132 Milano, Italy

THERE are several demonstrations that misfolded or unassembled proteins are not transported along the secretory pathway, but are retained intracellularly, generally in the endoplasmic reticulum (reviewed in ref. 1). For instance, B lymphocytes synthesize but do not secrete IgM, and only the polymeric form of IgM is secreted by plasma cells. The C-terminal cysteine of the μ heavy chain of secreted IgM (residue 575) is involved in the intracellular retention of unpolymerized IgM subunits². Here we report that the addition of reducing agents to the culture medium, at concentrations which do not affect cell viability, terminal glycosylation, or retention of proteins in the endoplasmic reticulum through the KDEL mechanism³, induces secretion of IgM assembly intermediates by both B and plasma cells. Free joining (J) chains, which are not normally secreted by plasma cells unless as part of IgM or IgA⁴, are also secreted in the presence of reducing agents. We propose a role for free thiol groups in preventing the unhindered transport of proteins through the secretory pathway. Under the scheme, assembly intermediates interact through their thiol groups between themselves and/or with unknown proteins of the endoplasmic reticulum. Such interactions may be prevented by altering the intracellular redox potential or by site-directed mutagenesis of the relevant cysteine residue(s).

The effect of the reducing agent 2-mercaptoethanol (2-ME) is described in the following examples. I.29 μ^+ B lymphoma cells do not secrete IgM, although they produce μ heavy (H) chains of the secreted form (μ_s) and assemble them with λ light (L) chains^{2,5}. The post-translational block of IgM secretion is reversed when pulse-labelled I.29 μ^+ cells are cultured with 2-ME (Fig. 1). Secretion is probably underestimated in Fig. 1, as part of the intracellular μ chains are of the membrane form (μ_m)^{2,5} and, during the chase, there is some intracellular degradation. On SDS-PAGE, secreted molecules migrate more slowly than intracellular ones, indicating that addition of 2-ME does not block terminal glycosylation and therefore does not seem to distort the normal pathway of secretion.

A chimaeric IgG2b molecule, in which the complete γ 2b chain is extended at the C terminus with the tailpiece (the last 20 amino acids) of μ_s (γ 2b- μ tp), is not secreted by J558L myeloma transfectant cells². Secretion can be restored by the addition of 2-ME, which is almost as efficient as the restoration achieved by mutating the C terminal cysteine to serine (Fig. 2a). Secretion of chimaeric IgG containing serine is not influenced by 2-ME. Titration experiments show that the induced secretion begins to level off at 1.4 mM 2-ME. At this concentration, most secreted material retains hapten (nitrophenacyl, NP) binding activity, which requires H-L chain assembly⁶ (Fig. 2b).

The time at which 2-ME is added does not seem to be important. Even when added 2 h after initiating the chase, 2-ME remains almost as effective at inducing γ 2b μ tp secretion (Fig. 2c). (All values are relative to total synthesis at the end of the chase, and therefore do not take into consideration the degradation of labelled intracellular material during the 2 h before the

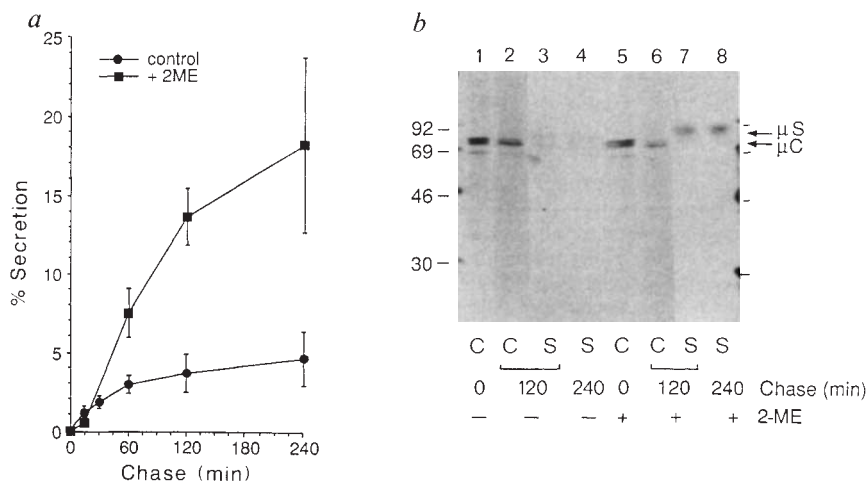
Received 4 May; accepted 20 July 1990.

- Ward, E. S., Güssow, D., Griffiths, A. D., Jones, P. T. & Winter, G. *Nature* **341**, 544–546 (1989).
- Better, M., Chang, C. P., Robinson, R. R. & Hurwitz, A. H. *Science* **240**, 1041–1043 (1988).
- Huston, J. S. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5879–5883 (1988).
- Bird, R. E. et al. *Science* **242**, 423–426 (1988).
- Chaudhary, V. K. et al. *Nature* **339**, 394–397 (1989).
- Glockshuber, R., Malia, M., Pfitzinger, I. & Plückthun, A. *Biochemistry* **29**, 1362–1367 (1990).
- Sastry, L. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5728–5732 (1989).
- Udada, K. et al. *Molec. Immun.* **27**, 25–35 (1990).
- Harper, M., Lema, F., Boulot, G. & Poljak, R. J. *Molec. Immun.* **24**, 97–108 (1987).
- Boulot, G. et al. *J. molec. Biol.* **213**, 617–619 (1990).
- Mariuzza, R. A. et al. *J. molec. Biol.* **170**, 1055–1058 (1983).
- Amit, A. G., Mariuzza, R. A., Phillips, S. E. V. & Poljak, R. J. *Science*, **233**, 747–753 (1986).
- Brünger, A., Kurian, J. & Karplus, M. *Science* **235**, 458–460 (1987).
- Luzzati, V. *Acta crystallogr.* **5**, 802–810 (1952).
- Colman, P. M. et al. *Nature* **326**, 358–363 (1987).
- Stevens, F. J., Chang, C. H. & Schiffer, M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6895–6899 (1988).
- Lascombe, M. B. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 607–611 (1989).
- Davies, D. R., Sheriff, S. & Padlan, E. A. *J. biol. Chem.* **263**, 10541–10544 (1988).
- Levitt, M. & Perutz, M. F. *J. molec. Biol.* **201**, 751–754 (1988).
- Kabsch, W. *J. appl. Crystallogr.* **21**, 916–924 (1988).
- Jones, A. *J. appl. Crystallogr.* **11**, 268–272 (1978).
- Plügrath, J. W., Saper, M. A. & Quiocho, F. A. in *Methods and Applications of Crystallographic Computing* (eds Hall, S. & Ashikawa, I.) 404–407 (Clarendon, Oxford, 1984).
- Sheriff, S., Hendrickson, W. A. & Smith, J. L. *J. molec. Biol.* **197**, 273–296 (1987).
- Tulp, W. et al. *Cold Spring Harbor Symp. quant. biol.* **54**, 257–263 (1989).

ACKNOWLEDGEMENTS. This work was supported by the CNRS, Institut Pasteur, Fondation Louis Jeantet and the US Public Health Service. We thank the CONVEX Company for access to a C210 computer.

§ To whom correspondence should be addressed.

FIG. 1 2-ME induces IgM secretion in B lymphomas. **a**, Induction of IgM secretion in B lymphoma cells. Kinetics of secretion of IgM by B lymphoma cells ($1.29\mu^+$) in the presence (squares) or absence (circles) of reducing agents. Data show average of three independent experiments $\pm$ s.e.m. $1.29\mu^+$ B lymphoma cells^{2,5} were pulsed for 15 min with ^{35}S methionine, washed and chased with or without 2-ME (7.1 mM), for the indicated times before immunoprecipitation of NP-40 (NP-40) cell lysates and supernatants with purified, isotype-specific anti- μ antibodies and protein A-Sepharose, analysis by SDS-PAGE and densitometric scanning of the autoradiogram. Two exposures of each autoradiogram were scanned. Samples were first pre-cleared with protein A-Sepharose alone. Immunoprecipitates were washed three times with 0.5 M NaCl, 50 mM Tris-HCl, pH 7.6, 0.25% NP-40, 1 mM phenylmethylsulphonyl fluoride, once in 5 mM Tris-HCl, pH 7.6, and eluted by boiling in SDS. The amount of μ chains secreted was calculated according to the formula: (area of secreted μ chain bands at times shown)/(area of intracellular μ chain (μ_c) band at the end of pulse ($t=0$)) $\times 100$. As part of the intracellular μ chain band is constituted by heavy chains of the membrane form (μ_m)^{2,5} IgM secretion is probably underestimated. The μ_c band decreased from 100% to 60% after a 240-min incubation. Addition of 2-ME during the pulse did not improve secretion. **b**, μ chains secreted in the presence of reducing agents are terminally glycosylated. Anti- μ immunoprecipitates from cell lysates (C, lanes 1, 2, 5–6) and



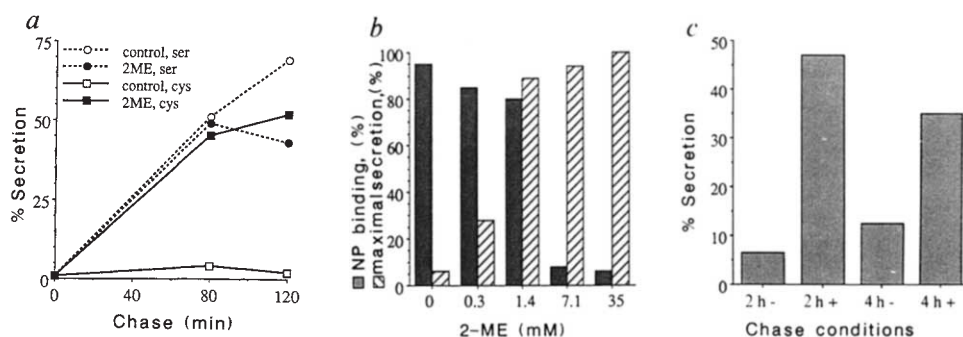
addition of 2-ME.) Thus the reason why the 2-ME induced secretion reaches a plateau after 1 or 2 h (Figs 1 and 2) is unlikely to be accounted for by molecules becoming irreversibly trapped intracellularly; oxidation of 2-ME over the incubation period may be more important.

J chains are thought to mediate IgM polymerization, forming disulphide bridges with Cys 575 (ref. 4). The excess J chain produced by myeloma cells (all of it in the absence of μ or α chains) is retained intracellularly and degraded⁷. But free J

supernatants (S, lanes 3, 4, 7, 8) of $1.29\mu^+$ cells chased in the presence (lanes 5–8) or absence (lanes 1–4) of 7.1 mM 2-ME for the indicated times. Lanes 1 and 5 show cell lysates at the end of the 15-min pulse. Marks on the left and right hand margins indicate relative molecular mass markers ($M_r \times 10^{-3}$) of two gels (lanes 1–4 and 5–8); arrows on the right hand margin indicate the mobility of intracellular and secreted μ (μ_c and μ_s , respectively). For a quantitative comparison, autoradiographs were exposed for 32 h (lanes 1, 2, 5, 6), and 75 h (lanes 3, 4, 7, 8).

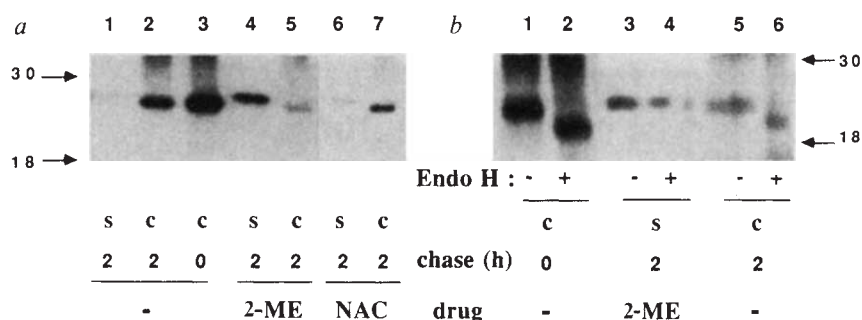
chains are secreted by myeloma cells incubated with 2-ME or *N*-acetyl cysteine (Fig. 3a). Secreted J chains migrate more slowly than the intracellular form. This is due to differential glycosylation as demonstrated by digestion with endoglycosidase H. Intracellular J chains are entirely sensitive to the enzyme, whereas the molecules secreted by cells treated with 2-ME are resistant (Fig. 3b). J chains are retained in a pre-Golgi compartment, probably the endoplasmic reticulum (ER)⁷; treatment with 2-ME presumably induces transit through the Golgi,

FIG. 2 Induced secretion of retained molecules expressed in plasma cells. **a**, Secretion of chimaeric IgG2b containing the μ_s tailpiece. Kinetics of secretion of IgG2b- μ tp (squares) and IgG2b- μ tp_{ser} (circles), in the presence (closed symbols) or absence (open symbols) of 1.4 mM 2-ME. J558L cells were transfected with a plasmid (pSV-V γ 2b- μ tp) encoding a γ 2b chain extended at the C terminus with the μ_s tailpiece (J[γ 2b- μ tp]), or with a mutant in which the C-terminal cysteine was replaced by serine (J[γ 2b- μ tp_{ser}])². J[γ 2b- μ tp_{ser}] and J[γ 2b- μ tp] cells were pulsed for 15 mins with ^{35}S methionine, and chased for 80 or 120 min in the presence or absence of 2-ME (1.4 mM) before precipitation of NP-40 cell lysates and supernatants with sheep anti-Ig antibodies and protein A-Sepharose or NP-Sepharose and SDS-PAGE under reducing conditions (12% acrylamide). Fluorograms of the gels were scanned by a densitometer, and the percentage of secreted immunoglobulin determined as described in legend to Fig. 1. Abbreviations: cys, J[γ 2b- μ tp]; ser, J[γ 2b- μ tp_{ser}]. The kinetics of free λ secretion were similar in the two cell lines (J[γ 2b- μ tp_{ser}] and J[γ 2b- μ tp]) and were not significantly affected by 2-ME (data not shown). **b**, Titration of 2-ME. Effects of different concentrations of 2-ME on the percentage of total secreted (hatched bars) or NP-binding chimaeric (stippled bars) IgG2b- μ tp in the supernatants of J[γ 2b- μ tp]. J[γ 2b- μ tp] cells were pulsed with ^{35}S methionine, and chased for 2 h in the presence of different concentrations of 2-ME, before precipitation of NP-40 cell lysates and supernatants



with rabbit anti-mouse Ig antibodies and protein A-Sepharose or Sepharose-bound NP and SDS-PAGE under reducing or nonreducing conditions. Samples were first pre-cleared with Sepharose beads coated with normal mouse serum. Densitometric analyses were performed on at least two exposures of each autoradiogram, and the percentage of secretion calculated as described in legend to Fig. 1. Data are expressed as a percentage of maximal secretion (hatched bars), which was reached at 35 mM 2-ME. The percentage of secreted material that retained hapten-binding activity (stippled bars) was calculated according to the formula $(100 \times \text{NP-precipitated radioactivity})/(\text{anti-Ig-precipitated radioactivity})$. **c**, 2-ME is effective long after protein synthesis. Pulse-labelled J[γ 2b- μ tp] were chased for 2 or 4 h in medium (2 h-; 4 h-). 2-ME was added either immediately (2 h+) or 2 h after initiation of the chase (4 h+). In both cases, cells were chased for 2 h in the presence of 2-ME (1.4 mM). Samples were immunoprecipitated, and autoradiograms of SDS-PAGE analysed by densitometry. The percentage of secreted γ 2b- μ tp was calculated as described in the legend to Fig. 1, relative to intracellular γ 2b- μ tp chains at $t=0$.

FIG. 3 Secretion of free immunoglobulin J chain in the presence of reducing agents. *a*, 2-ME induces secretion of free J chains. J558L myeloma cells, pulsed for 15 min with ^{35}S methionine, were chased for 2 h in medium (lanes 1, 2), in the presence of 7.1 mM 2-ME (lanes 4, 5), or 1 mM *N*-acetyl cysteine (NAC, lanes 6, 7) before precipitation of NP-40 cell lysates (C, lanes 2, 3, 5, 7) and supernatants (S, lanes 1, 4, 6) with rabbit anti-mouse J chain antibodies (gift from R. M. E. Parkhouse) and protein A-Sepharose and SDS-PAGE under reducing conditions. Lane 3: cell lysates, no chase. Samples were first precleared with protein A-Sepharose alone. Control immunoprecipitates included no antibody or rabbit anti- λ chain (not shown). Note the shift in M_r between intracellular and extracellular J chains. Aliquots of the anti-J immunoprecipitates were treated with endoglycosidase H (see *b*). *b*, Acquisition of endoglycosidase H resistance by J chain secreted in the presence of 2-ME. After extensive washing, aliquots of the anti-J immunoprecipitates from J558L cells chased for 0 or 2 h in the presence or absence of 2-ME (C, cell lysates; S, super-



because it allows terminal glycosylation and exocytosis.

The low background amount of intracellular-form μ chains, and the virtual absence of cytosolic proteins such as actin or lactate dehydrogenase⁸ from the supernatants (Table 1), exclude the possibility that the effects of 2-ME are due to cell lysis. Furthermore, the immunoglobulin heavy chain-binding protein BiP, which is retained in the ER through a different mechanism, namely the C-terminal sequence KDEL (single-letter amino-acid code)³, is not secreted even at 7.2 mM 2-ME. Not all protein intermediates that are retained intracellularly can be exported in the presence of 2-ME. Free μ chains (with or without C-terminal cysteine), which are retained through interactions between BiP and the CH1 domain^{2,9,10}, are not secreted in the presence or absence of 2-ME. By contrast, secretion of mutant μ chains, in which the C μ 1 had been removed, can be induced by reducing agents (Table 1). These results indicate that disulphide interchange reactions are responsible for the intracellular retention of partially assembled intermediates. Addition of reducing agents is likely to disrupt the normal redox potential

natants; see above for details), were resuspended in 75 mM sodium citrate, pH 5.5, 50 mM 2-ME, 0.1% Triton X-100, 1 mM phenylmethylsulphonyl fluoride and treated with (lanes 2, 4, 6) or without (lanes 1, 3, 5) 1 mU endoglycosidase H (Boehringer).

gradient of the secretory pathway and permit secretion of retained proteins.

The results may also throw some light on the hitherto unexplained failure of B cells to polymerize and secrete IgM. When B lymphocytes differentiate into antibody-secreting plasma cells, there are large changes in the ER, together with a much increased rate of synthesis of heavy and light chains^{11,14} and of J chains³. These changes could include a modification of the redox gradient along the secretory pathway, which would favour both polymerization and transit through the ER. It remains to be seen whether such changes in the redox potential are correlated to the fact that B and T cells, but not plasma cells, require 2-ME for optimal growth^{12,13}, at a concentration just below the threshold necessary to induce secretion of retained intermediates.

The effects of reducing agents are likely to have wider implications than the ones described here for B and plasma cells. Addition of reducing agents might induce (or increase) secretion or membrane expression of other proteins where cysteine residues are involved in their intracellular retention. These might include unassembled, and in some cases misfolded proteins, and hence have relevant applications in biotechnology. Preliminary data suggest that reducing agents induce secretion of recombinant *Escherichia coli* β -galactosidase by yeast transformants (M. Vanoni *et al.*, personal communication), and upregulate surface expression of the transferrin receptor in T cell blasts (C. A. *et al.*, unpublished data). Indeed, the effects of reducing agents might be an indicator of whether the intracellular transport of a protein is controlled by —SH group-mediated interactions. □

TABLE 1 Selectivity of the effects of reducing agents

Intracellular protein	Cell type	Secretion*	
		-2-ME	+2-ME
IgM	B cell	-	+
λ chain	Plasma cell	+++	+++
J chain	Plasma cell	-	+
IgG2b μ tp [†]	Plasma cell	-	++
IgG2b μ tp _{ser} [†]	Plasma cell	++	++
Free μ chain [†]	Plasma cell (no L chain)	-	-
μ_{Ala} [†]	Plasma cell (no L chain)	-	-
$\mu\Delta$ CH1 [†]	Plasma cell (no L chain)	-	+
$\mu_{Ala}\Delta$ CH1 [†]	Plasma cell (no L chain)	+	+
BiP	Both B and plasma cells	-	-
LDH [‡]	Both B and plasma cells	-	-
Actin [§]	Both B and plasma cells	-	-

* Secretion was monitored by endogenous labelling, ELISA² or enzyme activity (lactate dehydrogenase, LDH)⁸.

[†] The γ 2b heavy chain of IgG2b μ tp has been extended at the C terminus with the last 20 residues of secretory μ chains, including Cys 575. In IgG2b μ tp_{ser}, the latter has been replaced by serine. μ_{Ala} is a mutant of the secretory μ chain in which cysteine 575 has been substituted with alanine. In $\mu\Delta$ CH1 the first constant domain has been deleted, with Cys 575 replaced by alanine in $\mu_{Ala}\Delta$ CH1. All the mutants and transfectants are described in detail elsewhere².

[‡] LDH activity was measured as described in Bergenmeyer *et al.*⁸.

[§] The presence of actin in tissue culture supernatants was determined by SDS-PAGE.

Received 25 May; accepted 6 August 1990.

- Hurtley, S. M. & Helenius, A. *Rev. Cell Biol.* **5**, 277-307 (1989).
- Sitja, R. *et al. Cell* **60**, 781-790 (1990).
- Munro, S. & Pelham, H. R. B. *Cell* **48**, 899-907 (1987).
- Koshland, M. E. *A. Rev. Immun.* **3**, 425-453 (1985).
- Sitja, R. *et al. Molec. Immun.* **25**, 189-197 (1988).
- Reth, M., Imanishi-Kari, T. & Rajewsky, K. *Eur. J. Immun.* **9**, 1004-1013 (1976).
- Mosmann, T. R., Gravel, Y., Williamson, A. R. & Bauman, R. *Eur. J. Immun.* **8**, 94-101 (1978).
- Bergenmeyer, H. U., Grassi, M. & Walter, H. E. In *Methods of Enzymatic Analysis* Vol. 2, 3rd edn (ed. Bergmeyer, H. U.) 232-233 (Verlag Chemie, Weinheim, 1983).
- Bole, D. G., Hendershot, L. M. & Kearney, J. F. *J. Cell Biol.* **102**, 1558-1566 (1986).
- Hendershot, L. M., Bole, D. G., Köhler, G. & Kearney, J. F. *J. Cell Biol.* **104**, 761-767 (1987).
- Cooper, M. D. & Burrows, P. in *Immunoglobulin Genes* (eds Honjo, T., Alt, F. W. & Rabbits, T. H.) 1-22 (Academic, London, 1989).
- Broome, J. D. & Jeng, M. W. *J. exp. Med.* **138**, 574-592 (1973).
- Opitz, H. G., Lemke, H. & Hewlett, G. *Immun. Rev.* **40**, 53-77 (1978).
- Wist, D. L. *et al. J. Cell. Biol.* **110**, 1501-1511 (1990).

ACKNOWLEDGEMENTS. We thank T. Creighton, M. S. Neuberger and M. Rocco for discussions, suggestions and reagents, R. M. E. Parkhouse for antibodies, A. Albertini and G. B. Ferrara for support and encouragement, D. Finazzi, A. Fra and C. Valetti for helping in many experiments and sharing reagents, and J. Whybrow for secretarial assistance. This work was supported in part by the Associazione Italiana Ricerca sul Cancro (R. S.).

Isolation and characterization of the gene for a yeast mitochondrial import receptor

Hiroshi Murakami, Günter Blobel & Debkumar Pain

Laboratory of Cell Biology, Howard Hughes Medical Institute, The Rockefeller University, New York, New York 10021, USA

WE have previously identified an integral membrane protein (p32) from *Saccharomyces cerevisiae* as a receptor for protein import into mitochondria, and have localized it to the mitochondrial outer membrane at contact sites¹. Here we report isolation of the corresponding mitochondrial import receptor gene, termed *MIR1*. The deduced amino-acid sequence of p32 shows roughly 40% identity with proteins of bovine heart² and rat liver³ that have been suggested to be mitochondrial phosphate carriers. Haploid cells carrying a disrupted *MIR1* allele were unable to grow on a non-fermentable carbon source but grew in media containing glucose, indicating that the *MIR1* protein is essential for mitochondrial function. Compared with wild type, amounts of some mitochondrial proteins were markedly reduced in cells containing a disrupted *MIR1* allele, whereas levels of others were unchanged. This indicates that yeast contains more than one pathway for protein import into mitochondria.

To synthesize a DNA probe for isolating the *MIR1* gene, we determined the N-terminal sequence of 27 residues of purified p32 (Fig. 1a). Partially degenerate sense and antisense oligonucleotides were synthesized and used as primers in a polymerase chain reaction (PCR) (Fig. 1b). Three independent clones harbouring a PCR product of the appropriate size were sequenced and found to encode the N-terminal amino-acid sequence of p32.

The *KspI-SalI* fragment of one of the clones (pNYHM5) (Fig. 1c) hybridized to a single size messenger RNA of 1.4 kilobases (kb) in a northern blot of total yeast RNA, suggesting that there is only one species of *MIR1*-derived mRNA (Fig. 2). There was more than five times as much *MIR1* mRNA in yeast grown on lactate (Fig. 2a, lane 1) compared with yeast grown on glucose (Fig. 2a, lane 2), whereas actin mRNA, serving as control, was present in equal amounts under both conditions

(Fig. 2b, lanes 1 and 2). The results of a Southern blot of yeast genomic DNA probed with the ³²P-labelled *KspI-SalI* fragment of pNYHM5 (Fig. 2c) suggested that there is only one copy of the *MIR1* gene in the haploid yeast genome.

We screened 3.5×10^5 plaques of a yeast genomic λ gt11 library with the ³²P-labelled *KspI-SalI* fragment of pNYHM5 and obtained five overlapping clones. One clone (λ NHM10) contained the entire coding region of the *MIR1* gene, and was used in all subsequent analyses. DNA sequence analysis of the λ NHM10 insert revealed a 933-base-pair (bp) open reading frame (ORF) that started with a methionine codon and was not interrupted by any yeast splicing consensus sequences (that is, 5' splice site consensus sequence GTA(T/C) GT, or branch point consensus sequence TACTAAC⁴). It encoded a polypeptide of 311 amino acids with a calculated relative molecular mass of 32,800 (M_r 32.8K). This value agrees well with the M_r estimated for p32 from its electrophoretic mobility in SDS-PAGE¹. The N-terminal amino-acid sequence deduced from *MIR1* ORF is identical to that determined for the first 27 amino acids of p32 except for the second residue (Ser instead of Tyr) (Figs 1a and 3a). This discrepancy may be due to differences between the yeast strains used for construction of the λ gt11 library and for purification of the p32 protein. Assignment of the methionine codon at position +1 (Fig. 3a) as the initiation codon is based on several criteria: (1) the identity of the deduced and determined N-terminal amino-acid sequences for p32; (2) the absence of any yeast splicing consensus sequences⁴ in the sequence shown in Fig. 3a, which could have disrupted the ORF (3) the lack of another in frame methionine codon between this methionine codon and the stop codon 16 codons upstream; and (4) the similarity of the initiation codon ATG context sequence of the *MIR1* gene, CTCACCATGTCT, with the corresponding yeast consensus sequence, (A/T)A(A/C)A(A/C)AATGTC-(T/C)⁵ (9 out of 12 nucleotides identical).

The initiation of p32 translation at methionine +1 indicates that p32 is not synthesized as a larger precursor with a cleavable signal sequence. This is supported by observations that the *MIR1* gene product synthesized *in vitro* showed an electrophoretic mobility on SDS-PAGE identical to that of endogenous p32, and that the *in vitro* synthesized product was imported into mitochondria without any detectable change in its M_r as determined by SDS-PAGE (data not shown). In addition, imported p32 was resistant to alkali extraction (data not shown), indicat-

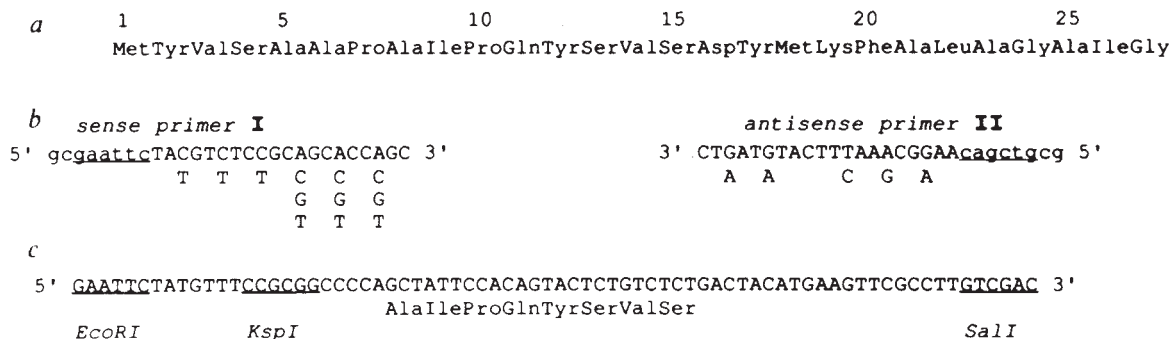


FIG. 1 Amino-terminal amino-acid sequence of a mitochondrial import receptor (p32) and synthesis of a p32-specific probe by PCR. **a**, N-terminal amino-acid sequence of purified p32; **b**, partially degenerate sense (I) and antisense (II) primers, lower case letters indicating 5' extension carrying *EcoRI* or *SalI* sites (underlined); **c**, nucleotide sequence of one of the PCR products; the *KspI-SalI* fragment of this clone (pNYHM5) was used as a probe for further studies.

METHODS. Protein p32 was purified¹, transferred to polyvinylidene difluoride membrane (Millipore) and sequenced on a gas-phase sequenator. Partially degenerate sense (I) and antisense (II) primers were synthesized on the basis of the amino-acid sequence obtained (a) and on codon usage analysis of *S. cerevisiae*¹², with additional 5' extensions carrying *EcoRI* and *SalI*

sites (b). Primers were further purified by Oligonucleotide Purification Cartridges (Applied Biosystems). The PCR¹³ (0.1 ml) contained 10 μ M of each primer, 1 ng ml⁻¹ of *S. cerevisiae* genomic DNA, 0.4 mM of each dNTP, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂, 0.01% gelatin and 2.5 units of *Taq* polymerase (Perkin Elmer Cetus). The initial cycle consisted of denaturation at 94 °C for 5 min, annealing at 37 °C for 2 min and extension at 72 °C for 3 min. The subsequent 39 cycles were carried out under identical conditions except that denaturation was for 1 min. Reaction products with expected length (~80 bp) were isolated and subcloned into *EcoRI-SalI* sites of pBluescript II SK(-) (Stratagene) and sequenced¹⁴ using Sequenase version 2 (US Biochemicals). The *KspI-SalI* fragment of one of the clones shown in Fig. 1c (pNYHM5) was used as a probe for further studies.

ing that it was integrated into the mitochondrial membrane without the aid of a cleavable signal sequence.

The isoelectric point of the *MIR1* gene product was calculated as 9.35. Its hydropathy profile showed six hydrophobic domains (Fig. 3b). Although only one of these (from residues 215–233) gave a sufficiently high hydrophobic index to qualify as a transmembrane segment⁶, some or all of the remaining five hydrophobic regions may serve as transmembrane segments.

Comparison of p32 with protein sequences in the National Biomedical Research Foundation Protein Identification Resources protein sequence database revealed 40% identity of p32 with proteins of bovine heart² (Fig. 3c) and rat liver³ that have been proposed to be the mitochondrial phosphate carrier. Identity is primarily observed in the hydrophobic regions.

To analyse the *in vivo* function of the *MIR1* gene product we generated a diploid heterozygous for a disrupted allele of *MIR1* (*mir1*⁻) and examined the fate of its haploid spores. All spores that contained a disrupted *MIR1* allele could still grow on glucose but could no longer grow on glycerol (petite phenotype) (data not shown), indicating disrupted mitochondrial function. As mitochondria carry out critical reactions in amino acid and lipid metabolism requiring import of the corresponding enzymes, complete disruption of import would be expected to be lethal. Because the *mir1*⁻ mutant grows on glucose, we conclude that the *MIR1* protein cannot be the only mitochondrial import receptor in yeast.

As the *MIR1* disruption did not affect viability on glucose, we compared the amounts of several mitochondrial proteins in cell lysates from wild-type and *mir1*⁻ cells by immunoblotting (Fig. 4a). As expected, *mir1*⁻ cells did not contain detectable amounts of the 32K *MIR1* gene product, whereas the amounts of other mitochondrial proteins varied from equivalent to much reduced. The mutant cells contained an equivalent amount of the outer membrane protein porin and of the matrix protein delta¹-pyrroline-5-carboxylate dehydrogenase. The amounts of F1-ATPase γ subunit (data not shown) and of cytochrome *c* oxidase subunit IV, both associated with the inner membrane were reduced by 74% and 73%, respectively. Cytochrome *c* oxidase subunit II, encoded by the *COXII* gene of mitochondrial DNA, was reduced by more than 94%. Identical results were obtained with isolated mitochondria from both wild-type and mutant strains (data not shown). These data further substantiate the existence of alternative pathway(s) for protein import into mitochondria.

As subunit II of cytochrome *c* oxidase is encoded by the *COXII* gene in mitochondrial DNA and as nuclear petite mutants tend to lose mitochondrial DNA, we examined whether the *mir1*⁻ mutant still has mitochondrial DNA. Yeast cells with a disrupted *MIR1* gene were transformed with a plasmid carrying both the intact *MIR1* gene and a *LEU2*⁺ marker. All 126 Leu⁺ transformants examined were able to grow on glycerol (data not shown), indicating that the petite phenotype and the significant reduction of *COXII* gene product of the *mir1*⁻ mutant were not due to loss of the mitochondrial genome. These data suggest that the *mir1*⁻ mutation could reflect an impaired import of protein(s) which are involved in mitochondrial protein synthesis and/or assembly of cytochrome *c* oxidase.

We isolated mitochondria of *mir1*⁻ and wild-type cells and examined them by electron microscopy. Compared with mitochondria of wild-type cells, those of *mir1*⁻ cells had underdeveloped cristae and often contained dense bodies in their matrix (Fig. 4b). The underdeveloped cristae are consistent with the reduced quantities of integral or peripheral proteins of the inner membrane in these cells.

We also tested isolated mitochondria of the *mir1*⁻ mutant for their ability to import proteins. The rate of protein import was reduced by 50–60% compared to wild-type mitochondria (data not shown). This reduction may result directly from the *MIR1* disruption, and/or it may be a result of reduced mitochondrial energy production, for example as a consequence of reduced

cytochrome *c* oxidase subunit II (Fig. 4a) and/or other proteins.

The amino-acid sequence of the yeast *MIR1* gene product is 40% identical to that of mammalian mitochondrial proteins^{2,3} tentatively identified as phosphate carriers. The yeast protein could be either the functional homologue of the two mammalian proteins or a functionally distinct but related member of a larger family of proteins.

Consistent with the notion that the yeast and mammalian proteins are functionally distinct is the fact that the yeast protein is not synthesized with a cleavable signal sequence (Fig. 3a) whereas the two mammalian proteins^{2,3} are (Fig. 3c). To our knowledge there is no precedent for the loss (or acquisition) in evolution of a signal sequence for structurally and functionally homologous proteins imported from the cytoplasm into mitochondria. Furthermore, although the location of the two mammalian proteins has not yet been determined, they should be located exclusively in the inner mitochondrial membrane if they are indeed phosphate carriers. In contrast, p32, the yeast *MIR1* protein, has been localized to the outer membrane at the contact sites by immunoelectron microscopy and subfraction-

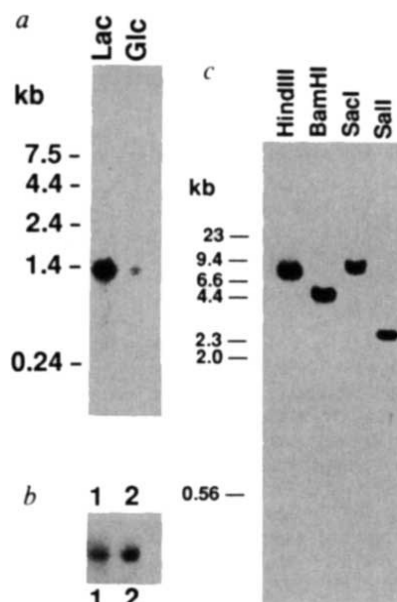


FIG. 2 Northern and Southern blot analyses. *a*, Northern blot of total RNA prepared from yeast grown either on lactate (lane 1) or on glucose (lane 2) probed with the *KspI*-*SalI* fragment of pNYHM5. Migration positions of RNA markers (BRL) are indicated on the left. *b*, Same blot probed with [³²P]-labelled chicken β -actin complementary DNA. *c*, Southern blot of yeast chromosomal DNA probed as in *a*. Restriction enzymes used are shown at the top of each lane. Migration positions of DNA markers (*HindIII* digest of DNA lambda) are shown on the left.

METHODS. The *KspI*-*SalI* fragment of pNYHM5 was labelled by the primer extension reaction with primer II (Fig. 1b) and [³²P]dCTP under standard conditions¹⁵. A DNA fragment coding for chicken β -actin was radiolabelled with [³²P]dCTP by the random primed DNA labelling procedure¹⁶. Total RNA was isolated by guanidium thiocyanate extraction¹⁷ from yeast grown either in 2% lactate medium¹⁸ (lane 1) or in YPD medium¹⁹ (1% yeast extract, 2% peptone, 2% dextrose) (lane 2). RNA (20 μ g) was electrophoresed in a 1.2% agarose/0.37 M formamide gel and subsequently transferred to a nitrocellulose filter using standard procedures¹⁵. The RNA was hybridized with either the [³²P]-labelled *KspI*-*SalI* fragment of pNYHM5 or [³²P]-labelled β -actin DNA at 42 °C for 12 h in a solution of 30% formamide, 5 $\times$ SSC, 50 mM sodium phosphate buffer pH7.4, 5 $\times$ Denhardt's solution, 0.1% SDS and 0.1 mg ml⁻¹ denatured herring sperm DNA¹⁵. Restriction enzyme digests of *S. cerevisiae* genomic DNA (10 μ g each) were separated by agarose gel electrophoresis and transferred to Hybond-N nylon membrane (Amersham) in an alkaline solution (0.25 M NaOH and 1.5 M NaCl), and the DNA was crosslinked to the membrane by ultraviolet irradiation as described by the supplier. Hybridization with ³²P-labelled *KspI*-*SalI* fragment of pNYHM5 was carried out under the conditions described above.

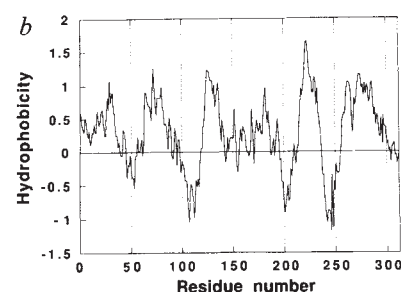
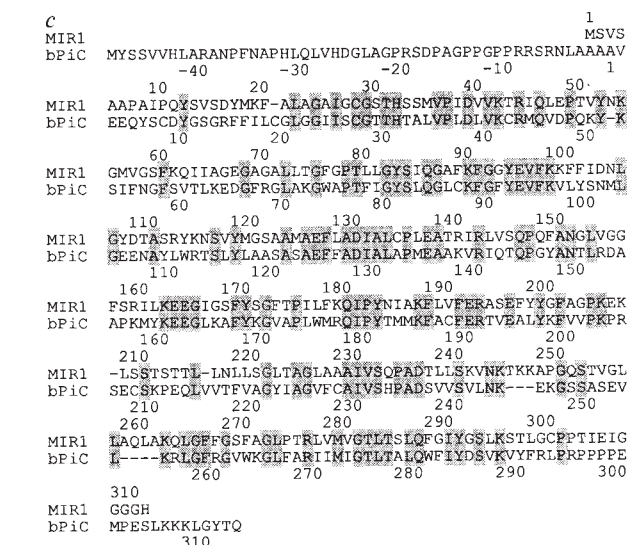


FIG. 3 Nucleotide sequence of the *MIR1* gene and analysis of the deduced amino-acid sequence. **a**, Nucleotide sequence of the DNA fragment containing the *MIR1* gene and the deduced amino-acid sequence. The first nucleotide and amino-acid residue of the translation start site are designated as position 1, and nucleotide and amino acid numbers are shown on the left of each row. The residues matching the N-terminal amino-acid sequence of purified p32 determined by Edman degradation (Fig. 1a) are indicated with double underlines. A TATAA-like sequence at positions -366 to -355, a polyadenylation signal at nucleotides 1,063 to 1,072, and *SalI*, *EcoRI* and *KpnI* sites are underlined. The stop codon is indicated by an asterisk. **b**, hydropathy profile of the *MIR1* gene product. The average hydrophobicity values were calculated⁶ with windows of 19 amino-acid residues. **c**, Alignment of the amino-acid sequences of the yeast mitochondrial import receptor (*MIR1*) and the putative phosphate carrier of bovine heart mitochondria (bPIC). The regions of similarity were identified²⁰ and sequences were aligned²¹ using the AALIGN program of DNASTAR software (DNASTAR Inc., Madison, WI) on an IBM AT computer. Identical residues are shadowed. Residue numbers of the *MIR1* protein and of bPIC are shown at the top and at the bottom, respectively. The first amino-acid residue of the mature form of bPIC is indicated by +1.

METHODS. A *S. cerevisiae* genomic λ gt11 library was screened with the [³²P]-labelled *KspI*-*SalI* fragment of pNYHM5 by standard procedures¹⁵ using the hybridization condition described in Fig. 2. Positive clones were plaque-purified, and their DNA was isolated and analysed by restriction enzyme digestion. The *SacI* fragment derived from one of the positive clones (λ NHM10) containing the entire *MIR1* gene was subcloned into pUC118 (ref. 22) in both orientations. The resulting clones (pNYHM77 and pNYHM78)

ation of mitochondrial membranes¹, a location also supported by the ability of anti-p32 antibodies to aggregate mitochondria from suspension and to inhibit the import of several proteins into mitochondria¹.

The rat liver and bovine heart proteins might, however, function as import receptors rather than phosphate carriers. The assignment of these proteins as phosphate carriers is not definitive as their purification to homogeneity and reconstitution into proteoliposomes capable of phosphate transport has not yet been accomplished⁷.



were further characterized. A series of nested deletions on these plasmids was generated by digestions with either exonuclease III and mung bean nuclease²³ or appropriate restriction enzymes. Single-stranded DNA was prepared as described²² and sequenced by the dideoxy chain termination method¹⁴ using Sequenase version 2 (US Biochemicals). Both strands were sequenced at least once.

Alternatively, the *MIR1* protein and its mammalian homologues could be bifunctional, serving as mitochondrial import receptor and phosphate carrier. A subunit (PEP) of the mitochondrial matrix signal peptidase also functions as a subunit of the cytochrome reductase complex of the respiratory chain⁸, serving as a precedent for the existence of bifunctional proteins in mitochondria.

Our data also suggest the existence of more than one receptor for protein import into mitochondria. This notion is consistent with the difficulties encountered by several laboratories in iden-

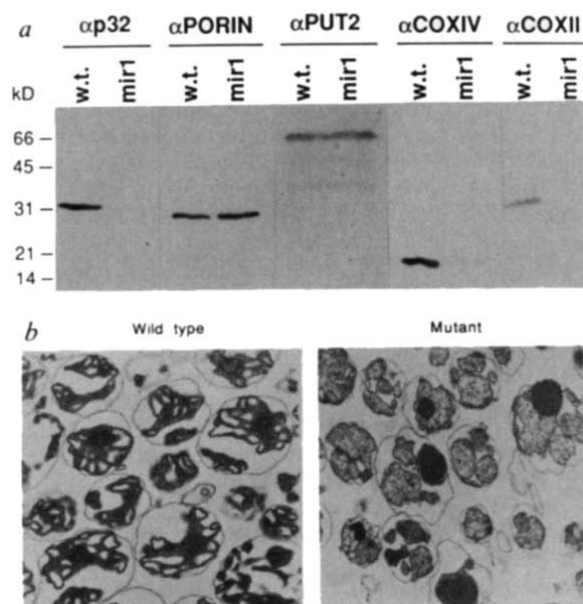


FIG. 4 Immunoblots of yeast lysates from wild-type and *mir1*⁻ cells and morphology of their mitochondria. *a*, Immunoblots of yeast lysates from wild type (w.t.) and *mir1*⁻ (*mir1*) cells. *M_r* markers are shown on the left. *b*, Electron micrographs of mitochondria isolated from wild-type and *mir1*⁻ mutant cells.

METHODS. Disruption of the *MIR1* allele. Plasmid pNYHM93 (a derivative of pNYHM78) carrying a 2.8-kb insert containing the *MIR1* gene, was digested with *Sall*, followed by treatment with the Klenow fragment of DNA polymerase I to generate blunt ends. A partial digestion of the *Sall* fragment with *EcoRI* yielded a DNA fragment that lacked the *Sall*-*EcoRI* fragment of the *MIR* coding region (Fig. 3*a*). This fragment was ligated to a 1.1-kb *SmaI*-*EcoRI* fragment carrying the URA3 gene derived from the plasmid YEp24. The 3-kb *SacI*-*SphI* fragment of the resulting plasmid pNYHM96, which harbours the *MIR1* gene disrupted by the URA3 gene, was used to transform the recipient yeast strain CGDi (heterozygote of CG378 and CG279 provided by Yeast Genetic Stock Center, Berkeley, CA) by one-step disruption²⁴. The resulting Ura3⁺ transformants were colony-purified and the disruption of one of the *MIR1* alleles was confirmed by Southern blot analysis of their chromosomal DNA using the *Sall*-*KpnI* fragment of the *MIR1* gene (Fig. 3*a*) as a probe (data not shown). Sporulation and tetrad analysis were performed using standard methods¹⁹. Preparation of yeast extracts and immunoblotting. Yeasts (CG389 as a wild-type strain and one of the *MIR1* disrupted strains) were cultured in YPD medium¹⁹ to mid-log phase and collected by centrifugation at 1,800g for 5 min. Cells were suspended in a solution of 0.5 M Trizma base, 6.25% SDS, 0.1 M DTT and 2 mM phenylmethyl sulphonyl fluoride and incubated in a boiling water bath for 5 min. Cell debris was removed by centrifugation in an Eppendorf bench top centrifuge at 10,000g for 10 min. The supernatants were collected and equivalent amounts of protein were separated by SDS-PAGE and analysed by immunoblotting with antisera against p32, porin, delta²-pyrroline-5-carboxylate dehydrogenase (PUT2), cytochrome *c* oxidase subunit IV (COXIV) and subunit II (COXII). Mitochondria were isolated as described previously¹⁸ from cells grown in 1% yeast extract, 2% peptone and 2% galactose. Mitochondria were fixed in a Karnovsky's aldehyde fixative containing 0.6 M sorbitol for 12 h at 4 °C and subsequently pelleted by centrifugation at 10,000g for 10 min. After osmication for 1 h at 4 °C, pellets were stained with uranyl acetate, dehydrated and embedded in Epon/Araldite. Sections were post-stained with uranyl acetate and lead citrate and examined with a JOEL 100CX electron microscope at an acceleration voltage of 80 kV.

tifying components of the mitochondrial import machinery by genetic approaches⁹. The fact that yeast contains several genes for some proteins of similar or identical function^{10,11} supports this idea. □

- Hamilton, R., Watanabe, C. K. & de Boer, H. A. *Nucleic Acids Res.* **15**, 3581-3593 (1987).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105-132 (1982).
- Kaplan, R. S., Pratt, R. D. & Pedersen, P. L. *Meth. Enzym.* **173**, 732-745 (1989).
- Schulte, U. *et al. Nature* **339**, 147-149 (1989).
- Brandriss, M. C. in *Protein Transfer and Organelle Biogenesis* (eds Das, R. C. & Robbins, P. W.) 655-676 (Academic, San Diego, 1988).
- Werner-Washburne, M., Stone, D. E. & Craig, E. A. *Molec. cell. Biol.* **7**, 2568-2577 (1987).
- Lawson, J. E. & Douglas, M. G. *J. biol. Chem.* **263**, 14812-14818 (1988).
- Maruyama, T., Gojobori, T., Aota, S. & Ikemura, T. *Nucleic Acids Res.* **14**, r151-r197 (1986).
- Mullis, K. B. & Faloona, F. A. *Meth. Enzym.* **155**, 335-350 (1987).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).
- Feinberg, B. & MacLaughlin, C. S. in *Yeast: A Practical Approach* (eds Campbell, I. & Duffus, J. H.) 147-161 (IRL, Oxford 1988).
- Murakami, H., Pain, D. & Blobel, G. *J. Cell Biol.* **107**, 2051-2057 (1988).
- Sherman, F., Fink, G. R. & Hicks, J.B. *Laboratory Course Manual for Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, New York, 1986).
- Lipman, D. J. & Pearson, W. R. *Science* **227**, 1435-1441 (1985).
- Needleman, S. B. & Wunsch, C. D. *J. molec. Biol.* **48**, 443-453 (1970).
- Vieira, J. & Messing, J. *Meth. Enzym.* **153**, 3-11 (1987).
- Henikoff, S. *Gene* **28**, 351-359 (1984).
- Rothstein, R. J. *Meth. Enzym.* **101**, 202-211 (1983).

ACKNOWLEDGEMENTS. We thank R. Young (Whitehead Institute, Cambridge, Massachusetts) for his yeast λ gt11 library and M. Douglas (University of North Carolina, Chapel Hill) and K. Mihara (Kyushu University, Japan) for their generous gifts of antisera. We thank D. Atherton and the Rockefeller University/Howard Hughes Medical Institute Biopolymer Facility for protein sequencing and synthesis of oligonucleotides and H. Shio for assistance with electron microscopy. We also thank D. Schnell, G. Sheinberg, G. Greenburg, N. Chaudhary, J. Sukegawa, E. Bartnik, R. Henriquez, J. Aris, F. Cross, C. Nicchitta and N. Schülke for their help on PCR, molecular cloning and tetrad analysis, and for critical reading of the manuscript.

RNA polymerase II C-terminal repeat influences response to transcriptional enhancer signals

Charles Scafe*, David Chao*, John Lopes†, Jeanne P. Hirsch†, Susan Henry† & Richard A. Young*

* Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Department of Biology, Carnegie Mellon University, 4400 5th Avenue, Pittsburgh, Pennsylvania 15213, USA

THE large subunit of RNA polymerase II contains a highly conserved and essential heptapeptide repeat (Pro-Thr-Ser-Pro-Ser-Tyr-Ser) at its carboxy terminus¹⁻⁷. *Saccharomyces cerevisiae* cells are inviable if their RNA polymerase II large subunit genes encode fewer than 10 complete heptapeptide repeats; if they encode 10 to 12 complete repeats cells are temperature-sensitive and cold-sensitive, but 13 or more complete repeats will allow wild-type growth at all temperatures³. Cells containing C-terminal domains (CTDs) of 10 to 12 complete repeats are also inositol auxotrophs⁸. The phenotypes associated with these CTD mutations are not a consequence of an instability of the large subunit³; rather, they seem to reflect a functional deficiency of the mutant enzyme. We show here that partial deletion mutations in RNA polymerase II CTD affect the ability of the enzyme to respond to signals from upstream activating sequences in a subset of promoters in yeast. The number of heptapeptide repeats required for maximal response to signals from these sequences differs from one upstream activating sequence to another. One of the upstream elements that is sensitive to truncations of the CTD is the 17-base-pair site bound by the GAL4 transactivating factor.

We investigated the ability of RNA polymerase II with a truncated CTD to transcribe three inducible genes *in vivo*. The mutation *rpb1Δ104* produces a large RNA polymerase II subunit with a CTD containing 11 3/7 heptapeptide repeats, rather than the 27 repeats found in wild-type cells³. The levels of *INO1*, *GAL10* and *HIS4* transcripts were determined in wild-type and in *rpb1Δ104* mutant cells, both before and after induction (Fig. 1). We chose *INO1* because RNA polymerase II CTD mutants

Received 18 May; accepted 14 August 1990.

- Pain, D., Murakami, H. & Blobel, G. *Nature* **347**, 444-449 (1990).
- Runswick, M. J., Powell, S. J., Nyren, P. & Walker, J. E. *EMBO J.* **6**, 1367-1373 (1987).
- Ferreira, G. C., Pratt, R. D. & Pedersen, P. L. *J. biol. Chem.* **264**, 15628-15633 (1989).
- Green, M. R. *A. Rev. Genet.* **20**, 671-708 (1986).

are deficient in *INO1* protein, and *GAL10* and *HIS4* because their promoters and the associated transcription factors are well characterized⁹⁻¹¹.

The ability to induce the synthesis of some messenger RNAs is diminished in cells containing the CTD deletion mutation *rpb1Δ104* (Fig. 1). In mutant cells, *INO1* mRNA was induced to 10% of the level found in wild-type cells, *GAL10* mRNA to 40%, and *HIS4* mRNA was induced to the same extent as in wild-type cells. These results indicate that the RNA polymerase II CTD mutant is defective in its response to induction of mRNA synthesis at some but not all promoters, and furthermore, that the extent of this deficiency varies at different promoters.

We found that the ability of wild-type and mutant RNA polymerases to respond to the *INO1*, *GAL10* and *HIS4* upstream activating sequence (UAS) elements parallels the degree to which these polymerases respond to induction at the complete promoters (Fig. 2). Response to the *INO1*, *GAL10* and *HIS4* regulatory elements was measured using *CYC1-lacZ*

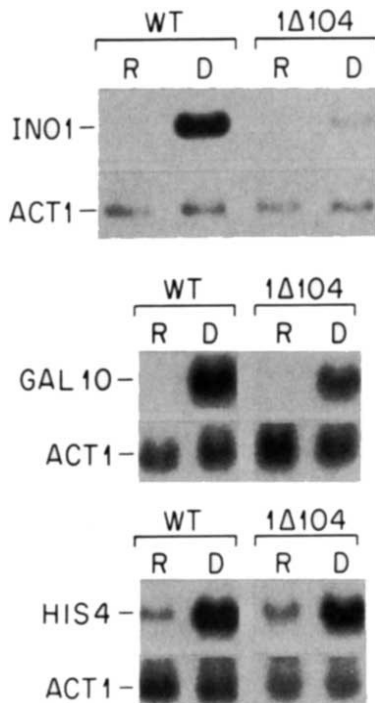


FIG. 1 Induction of *INO1*, *HIS4* and *GAL10* transcription in a CTD deletion mutant. The ability of RNA polymerase II from Z96 (wild type, WT) and mutant C6 (*rpb1Δ104*) cells to transcribe *INO1*, *HIS4* and *GAL10* was investigated before and after induction. RNA was prepared from these cells and an RNA gel blot was probed with radiolabelled *INO1*, *HIS4*, *GAL10* and *ACT1* probes. R, repressing conditions; D, derepressing conditions.

METHODS. *INO1* RNA gel blot: The yeast strains Z96 (WT) and C6 (*rpb1Δ104*) are described in Table 1 and in more detail in refs 3 and 22. The wild-type and mutant strains were grown to an absorbance at 600 nm of unity in minimal medium²³ supplemented with uracil and 200 μ M inositol. The cultures were washed twice and resuspended in the same medium containing 10 μ M inositol. This low concentration of inositol was adequate for *INO1* induction, and was preferred to no inositol, which can give rise to artefacts associated with cell death. At 0 and 10 h after induction, RNA was prepared²² from 10 ml samples of each culture for a northern blot which was probed with radiolabelled *INO1* and *ACT1* RNAs. *HIS4* and *GAL10* northern blots: *HIS4* transcription in the strains Z96 and C6 was induced by addition of 3-aminotriazole (to 10 mM) to cultures grown in YNB minimal medium (Difco). Cells were collected at 0 and 3 h after induction and their RNA isolated²². Induction of *GAL10* mRNA was accomplished by growing cells in SC medium³ with 2% raffinose, followed by a shift to SC medium plus 2% raffinose and 5% galactose. Cells were collected at 0 and 3 h after the shift to galactose-containing medium, and their RNA prepared for northern blots which were probed with radiolabelled *HIS4* RNA and radiolabelled *GAL10* and *ACT1* DNAs. Similar results were obtained 10 h and 24 h after induction. Relative signal strengths on the autoradiographs were determined by densitometry.

fusion UAS-reporter vectors. CTD mutant cells containing the *INO1* UAS-insert plasmid produced 10% of the β -galactosidase found in wild-type cells, whereas mutant cells with the *GAL10* UAS-insert plasmid produced 40%, and mutant cells carrying the *HIS4* UAS-insert plasmid made the same amount of β -galactosidase found in wild-type cells. Hence, the extent to which the wild-type and mutant RNA polymerases can respond to the three UAS elements in the *CYC1-lacZ* fusion assay precisely parallels their ability to use the intact promoter.

We investigated the effect of the length of RNA polymerase II CTD on the ability of the transcription apparatus to respond to *INO1*, *GAL10* and *HIS4* UAS elements by introducing the *CYC1-lacZ* fusion vectors with the *INO1*, *GAL10* or *HIS4* UASs into cells containing RNA polymerase II with different CTD lengths. Assay of β -galactosidase indicates that the three UASs differ in their sensitivity to CTD length (Fig. 3). We found a substantial reduction in β -galactosidase induced in cells containing *CYC1-lacZ* fusion vectors with the *INO1* or *GAL10* UASs as the CTD decreased from 27 to 11 complete heptapeptide repeats. For these two UAS elements there seems to be a threshold number of repeats above which the response to the UAS is at wild-type levels and below which the response falls off sharply. The slopes of the two curves below their respective thresholds are very similar. The response to the *HIS4* UAS seems not to be very sensitive to CTD length over the range tested.

The defective response of CTD mutants to the *GAL10* UAS can be reproduced with a UAS element consisting of only the 17-base-pair consensus GAL4 binding site (Fig. 3). The magnitude of the effect is comparable to that found with the normal *GAL10* UAS element, suggesting that diminished response to GAL4 binding sites accounts for the defective response to the entire UAS, and that GAL4 is involved in the defective response. These results are consistent with the observation that CTD alterations can affect the response to activation by GAL4 mutants³¹.

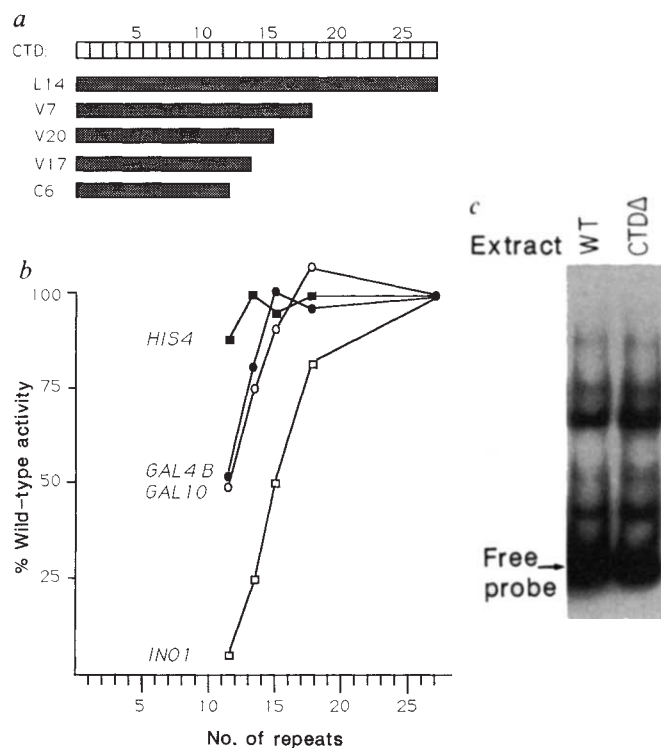
The inability of CTD mutant RNA polymerase to respond to signals from the *INO1* UAS is not due to differences in the levels of DNA-binding transcription factors (Fig. 3c). We find similar amounts of gel-shifted *INO1* promoter DNA with extracts from wild-type cells and CTD deletion mutant cells. Thus, the levels of DNA-binding transcription factors that associate with *INO1* promoter DNA are apparently not significantly affected by the CTD deletion, indicating that the lack of response to UAS-mediated signals is not due to any substantial change in the levels of *INO1* transcription factors. The reduced response to the *GAL10* UAS is probably not due to altered amounts of GAL4 protein; GAL4 mRNA levels are reduced twofold in CTD mutant cells, but the defect in the

TABLE 1 Yeast strains

Strain	Genotype
Z96	<i>MATα</i> , <i>his3Δ200</i> , <i>leu2-3,-112</i> , <i>ura3-52</i> , <i>rpb2Δ297::HIS3[pRP214(LEU2, RPB2)]</i>
C6	<i>MATα</i> , <i>his3Δ200</i> , <i>leu2-3,-112</i> , <i>ura3-52</i> , <i>rpb1Δ187::HIS3[pC6(LEU2, rpb1Δ104)]</i>
L14	<i>MATα</i> , <i>his3Δ200</i> , <i>leu2-3,-112</i> , <i>ura3-52</i> , <i>rpb1Δ187::HIS3[pL14(LEU2, rpb1Δ100)]</i>
V7	<i>MATα</i> , <i>his3Δ200</i> , <i>leu2-3,-112</i> , <i>ura3-52</i> , <i>rpb1Δ187::HIS3[pV7(LEU2, rpb1Δ111)]</i>
V17	<i>MATα</i> , <i>his3Δ200</i> , <i>leu2-3,-112</i> , <i>ura3-52</i> , <i>rpb1Δ187::HIS3[pV17(LEU2, rpb1Δ115)]</i>
V19	<i>MATα</i> , <i>his3Δ200</i> , <i>leu2-3,-112</i> , <i>ura3-52</i> , <i>rpb1Δ187::HIS3[pV19(LEU2, rpb1Δ117)]</i>
V20	<i>MATα</i> , <i>his3Δ200</i> , <i>leu2-3,-112</i> , <i>ura3-52</i> , <i>rpb1Δ187::HIS3[pV20(LEU2, rpb1Δ118)]</i>
N447	<i>MATα</i> , <i>his3Δ200</i> , <i>leu2-3,-112</i> , <i>ura3-52</i> , <i>rpb1Δ187::HIS3[pC3(LEU2, rpb1Δ103)]</i>
N460	<i>MATα</i> , <i>his3Δ200</i> , <i>leu2-3,-112</i> , <i>ura3-52</i> , <i>ade2</i> , <i>lys2Δ201</i>

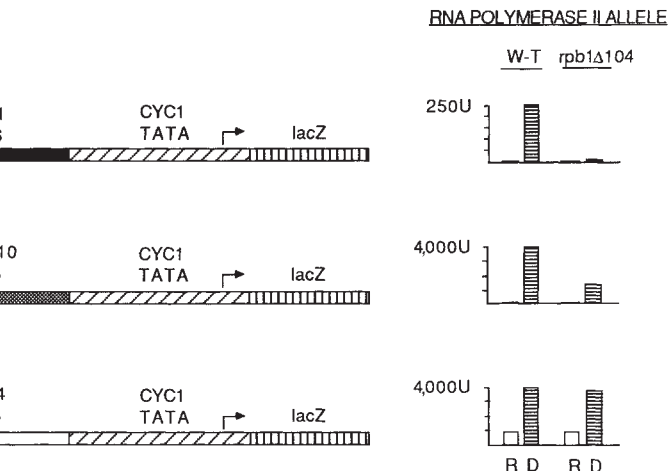
FIG. 2 The RNA polymerase II CTD partial deletion mutant is defective in utilization of some UAS elements. The relative ability of wild-type and *rpb1Δ104* RNA polymerase II to respond to *INO1*, *GAL10* and *HIS4* upstream activating sequences was measured using UAS-less *CYC1-lacZ* fusions and β -galactosidase assays. The three UAS elements have been placed in the same site relative to the *CYC1-lacZ* transcript start site. Plasmids pJH359, pLGSD5 and pHYC3 were maintained in the wild-type (Z96) and mutant (C6) strains as high copy 2 μ plasmids. Three independent transformants of each construct were grown under both repressing (R, open bar) and derepressing (D, bars with horizontal shading) conditions and their β -galactosidase activity measured. Units (U) are in nmol O-nitrophenyl- β -D-galactoside cleaved per min per mg protein.

METHODS. Plasmids pJH359, pHYC3 and pLGSD5 were introduced into the appropriate strains by lithium acetate transformation and maintained by selection for uracil prototrophy. The plasmid pJH359 contains a 205-base pair *Maell* DNA fragment encompassing the *INO1* UAS inserted at the *XhoI* site of the UAS-less *CYC1-lacZ* fusion vector pJH304. The plasmid pJH304 is a derivative of pLGΔ312 (ref. 24), in which the *CYC1* UAS elements have been removed. Plasmid pLGSD5 contains the *GAL10* UAS inserted at the



response to the *GAL10* UAS remains when *GAL4* mRNA levels are restored to wild-type by gene duplication.

We suggest that the RNA polymerase II CTD has a role in the dialogue between regulatory factors that bind upstream sites and general factors associated with DNA near the transcription start site in at least a subset of promoters. The observation that some promoters are more sensitive than others to CTD truncations *in vivo* is consistent with results obtained in transcription assays *in vitro*; purified RNA polymerase II lacking the heptapeptide domain cannot initiate transcription at the murine dihydrofolate reductase promoter¹², but is able to initiate transcription at several *Drosophila* promoters⁷ and the adenovirus-2 major late promoter^{12,13} *in vitro*. Precisely how the CTD functions is not clear, although direct interactions between the CTD



XhoI site to replace the *CYC1* UAS in the *CYC1-lacZ* fusion vector pLGΔ-292S (ref. 25). Plasmid pHYC3 contains the *HIS4* UAS inserted at the *XhoI* site to replace the *CYC1* UAS in the *CYC1-lacZ* fusion vector pLG669-Z (ref. 26). Inductions are described in Fig. 1 legend. Cell extracts for β -galactosidase assays were made by the glass-bead extraction procedure²⁷ and assayed according to Miller²⁸. Three independent transformants were tested for each strain, and the values given are the average of the three. Standard deviations were typically in the range of 5–20%.

FIG. 3 Response to UAS elements in a range of CTD deletion mutants. The ability of RNA polymerase with a variety of CTD repeat lengths to use *INO1*, *GAL10*, *HIS4* and *GAL4* binding site (*GAL4B*) UAS elements was measured using a *CYC1-lacZ* UAS-reporter vector and β -galactosidase assays. The plasmids pJH359, pLGSD5, pHYC3 and pSV14 were introduced into a series of strains containing *RPB1* CTDs of various lengths. Plasmid pSV14 contains a synthetic 17-base-pair *GAL4* binding site inserted into the UAS-less *CYC1-lacZ* fusion vector²⁹. **a**, The strains and their CTD repeat lengths are L14 (27 repeats), V7 (17 4/7), V20 (14 5/7), V17 (13 4/7) and C6 (11 3/7) (see Table 1 for genotypes). **b**, Results of β -galactosidase assays plotted as a percentage of activity obtained in extracts from the wild-type *RPB1* strain L14. Three independent transformants of each strain were exposed to inducing conditions and assayed for β -galactosidase. Standard deviations were typically in the range of 5–20%. **c**, Levels of DNA binding factors as measured in gel shift assays using an *INO1* UAS DNA probe and protein extracts from wild-type and mutant cells. The probe was a fragment containing *INO1* promoter DNA spanning the region from –259 to –155 relative to the start site of transcription. Yeast whole-cell extracts were prepared as described³⁰ from N460 (WT) and N447 (CTDΔ) grown in complete synthetic medium supplemented with 75 μ M inositol. Similar results were obtained with *rpb1Δ103* (N447) and *rpb1Δ104* (C6) mutants. The mobility shift assay has been described³⁰.

and transcription factors¹⁴, and between the CTD and DNA¹⁵ have been proposed.

The interaction between regulatory and general factors probably involves several mechanisms^{16–19} (reviewed in ref. 20). Our data indicate that one mechanism involves the highly conserved RNA polymerase II CTD, and it is possible that other components of the polymerase will have an accessory role. The recent isolation of genes for all ten yeast RNA polymerase II subunits²¹ should enable us to investigate the role of each of these conserved subunits in transcription. □

Received 23 May; accepted 6 August 1990.

- Allison, L. A., Moyle, M., Shales, M. & Ingles, C. J. *Cell* **42**, 599–610 (1985).
- Corden, J. L., Cadena, D. L., Ahearn, J. M. & Dahmus, M. E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7934–7938 (1985).
- Nonet, M., Sweetser, D. & Young, R. A. *Cell* **50**, 909–915 (1987).
- Ahearn, J. M. Jr, Bartolomei, M. S., West, M. L., Cisek, L. J. & Corden, J. L. *J. biol. Chem.* **262**, 10695–10705 (1987).
- Allison, L. A., Wong, J. K.-C., Fitzpatrick, V. D., Moyle M. & Ingles, C. J. *Molec. cell. Biol.* **8**, 321–329 (1988).
- Bartolomei, M. S., Halden, N. F., Cullen, C. R. & Corden, J. L. *Molec. cell. Biol.* **8**, 330–339 (1988).
- Zehring, W. A., Lee, J. M., Weeks, J. R., Jorke, R. S. & Greenleaf, A. L. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3698–3702 (1988).
- Nonet, M. & Young, R. A. *Genetics* **123**, 715–724 (1989).

9. Greenberg, M., Goldwasser, P. & Henry, S. A. *Molec. gen. Genet.* **186**, 157-163 (1982).
10. Guarente, L. *Cell* **52**, 303-305 (1988).
11. Struhl, K. A. *Rev. Biochem.* **58**, 1051-1077 (1989).
12. Thompson, N. E., Steinberg, T. H., Aronson D. B. & Burgess, R. R. *J. biol. Chem.* **264**, 11511-11520 (1989).
13. Kim, W.-Y. & Dahmus, M. *J. biol. Chem.* **264**, 3169-3176 (1989).
14. Sigler, P. *Nature* **333**, 210-212 (1988).
15. Suzuki, M. *Nature* **344**, 562-565 (1990).
16. Pugh, B. F. & Tjian, R. *Cell* **61**, 1187-1197 (1990).
17. Berger, S. L., Cress, W. D., Cress, A., Triesenberg, S. J. & Guarente, L. *Cell* **61**, 1199-1208 (1990).
18. Kelleher, R. J., Flanagan, P. M. & Kornberg, R. D. *Cell* **61**, 1209-1215 (1990).
19. Stringer, K. F., Ingles, C. J. & Greenblatt, J. *Nature* **345**, 783-786 (1990).
20. Lewin, B. *Cell* **61**, 1161-1164 (1990).
21. Woychik, N. & Young, R. A. *Trends biochem. Sci.* **15**, 347-351 (1990).
22. Scafe, C., Nonet, M. & Young, R. A. *Molec. cell. Biol.* **10**, 1010-1016 (1990).
23. Sherman, F., Fink, G. R. & Hicks, J. B. *Laboratory Course Manual for Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1986).
24. Guarente, L. & Mason, T. *Cell* **32**, 1279-1286 (1983).
25. Guarente, L., Yocum, R. R. & Gifford, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7410-7414 (1982).
26. Hinnebusch, A. G., Lucchini, G. & Fink, G. R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 498-502 (1985).
27. Rose, M., Casadaban, M. J. & Botstein, D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2460-2464 (1981).
28. Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1972).
29. Giniger, E., Varnum, S. M. & Ptashne, M. *Cell* **40**, 767-774 (1985).
30. Company, M., Adler, C. & Errede, B. *Molec. cell. Biol.* **8**, 2545-2554 (1988).
31. Allison, L. A. & Ingles, J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2549-2598 (1989).

ACKNOWLEDGEMENTS. C.S. and D.C. contributed equally to this work. We thank D. Baltimore, G. Fink, P. Sharp and R. Weinberg for discussion. This work was supported by grants from the Public Health Service (S.H. and R.Y.). C.S. was supported by a National Science Foundation predoctoral fellowship and J.L. by an NIH postdoctoral fellowship. R.A.Y. is a Burroughs Wellcome Scholar.

Identification of an actin-binding protein from *Dictyostelium* as elongation factor 1a

Fan Yang, Mark Demma, Vivien Warren, Suranganie Dharmawardhane & John Condeelis

Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, New York 10461, USA

INDIRECT evidence has implicated an interaction between the cytoskeleton and the protein synthetic machinery¹⁻⁸. Two recent reports have linked the elongation factor 1a (EF-1a) which is involved in protein synthesis, with the microtubular cytoskeleton⁷⁻⁸. *In situ* hybridization has, however, revealed that the messages for certain cytoskeletal proteins are preferentially associated with actin filaments⁶. ABP-50 is an abundant actin filament bundling protein of native relative molecular mass 50,000 (50K)⁹ isolated from *Dictyostelium discoideum*. Immunofluorescence studies show that ABP-50 is present in filopodia and other cortical regions that contain actin filament bundles^{9,10}. In addition, ABP-50 binds to monomeric actin in the cytosol of unstimulated cells and the association of ABP-50 with the actin cytoskeleton is regulated during chemotaxis¹⁰. Through complementary DNA sequencing and subsequent functional analysis, we have identified ABP-50 as *D. discoideum* EF-1a. The ability of EF-1a to bind reversibly to the actin cytoskeleton upon stimulation could provide a mechanism for spatially and temporally regulated protein synthesis in eukaryotic cells.

To investigate the mechanism of actin binding of ABP-50 and its regulation by chemotactic stimulation, ABP-50 was cloned and sequenced. Affinity purified polyclonal antibodies raised against ABP-50⁹ were used to screen a λ gt11 cDNA library (kindly provided by Peter Devreotes of Johns Hopkins University) made from the AX-3 strain of *D. discoideum* which had been allowed to progress to an early stage of the developmental cycle. Two cDNA clones, p50-1.7 (1.7 kilobase (kb)) and p50-1.5 (1.5 kb), were isolated and sequenced. The two clones largely overlapped one another, had complete amino-acid identity in the overlap region, and spanned the entire coding sequence of the protein (Fig. 1). The composite cDNA is 1.8 kb long, has 0.3 kb of 5' untranslated sequence, codes for a 456 amino-acid polypeptide with a calculated relative molecular mass of 50,090, and terminates in a poly(A) tail. The two cDNAs, which overlap

<i>a</i>			
ABP-50	MEFPESEKTHINIVVGHVDAGKSTTTGHLYKCGGIDKRVIEKYEKEAS	50	
Human	MGK.....S.....T.....F.....A		
Tomato	MGK..I..S.....L.....RF.....A		
Artemia	MGK..I..S.....S.....T.....F.....Q		
Yeast	MGK..S..V.....S.....T.....F.....A		
Mucor	MGK..V.V.....S.....T.....EF.....A		
	-----A-----		
ABP-50	EMGKQSFYAWMDKLKAERGITIDIALWKFETSKYVFTIIDAPGHRD	100	
Human	...G.....L.....S.....V.....		
Tomato	...N.R.....L.....T.....C.V.....		
Artemia	...G.....L.....A.....V.....		
Yeast	...L.G.....L.....P.QV.V.....		
Mucor	...L.G.....L.....P.NV.V.....		
	-----C-----		
ABP-50	FIKNMITGTSQADCAVLVIASTPGEFEAGIAKNGQTRHALLAYTLGVKQ	150	
Human	IV.AGV.....S.....		
Tomato	I.D.T.G.....S.D.....F.....		
Artemia	IV.AGV.....S.....		
Yeast	I.I.GGV.....S.D.....F.....R.		
Mucor	I.I.GG.....S.D.....F.....FR.		
	-----E-----		
ABP-50	MIVAINKMEKSTNYSQARYDEIVKEVSSFIKKIGYNPEKVAFPVIGWN	200	
Human	L..GV.....STEPP.....K.E..K.....TY.....DT.....		
Tomato	...CCC.....ATTPK.....K.....YL..V.....D.IP.....FE		
Artemia	L..GV.....STEPP.E..FE..K.....AY.....D.AA.....H		
Yeast	L.....V.....SV--KWDES.FQ.....T.N.....V.....KT.P.....		
Mucor	L.....TT--KW..D.N.....G.....F.KS.P.....H		
	-----G----- -----Actin binding?-----		
ABP-50	GDNMLERSDKMEWYK-----GPTLLEALDAIVEPKRPHDKPLR	238	
Human	P.AN.P.F.GWKVTRKDNAS.T.....C.LP.T.T.....		
Tomato	I.....TNLD.....Q.N.....S.....		
Artemia	A..RLP.....GWNIERKEGKAD.K..D.....LP.S..TE.....		
Yeast	I.ATTNAP.....GWEKETKAGVVK.K..D.I.....EQ.S..T.....		
Mucor	DE.TN.P.F.GWNKETKAGSKT.K.....I.....EP.V..S.....		
	-----J-----		
ABP-50	IPLQDVYKIGGIGTVFVGRVETGIKPGMVVTFAPAGLSTEVKSVEMHHE	288	
Human	VL.....V.....VNV.....		
Tomato	V.....G.T.T.....		
Artemia	L.....V.....I.....NIT.....		
Yeast	L.....V.....VT.....		
Mucor	L.....T.A.....N.....AVT.....		
	-----K-----		
ABP-50	QLPEARPGDNVGFNVKNSVKEIKRGMVAGDSKNDPPQETEFVAQVIVL	338	
Human	A.S..L.....DVR..N.....N..M.AAG.Y.....I.		
Tomato	A.Q..L.....A.DL..Y..SN..D..AKGAAS.T.....IM		
Artemia	S.EQ.S.....LR..Y..S.....N.ARGSD.F.....		
Yeast	..EQGV.....R..N.C..A.....KGCAS.N.T.....		
Mucor	T.T.GL.....D.R..N.CS.....AK.SAS.T.....I.		
ABP-50	NHPGQIHAGYSPVLDCHTAHIAKCFTEIVDKVDRRTGAVVAREGTAAVVL	388	
Human	S..A.....A.LKE.I.....KKLEDGPK---F.		
Tomato	GN..A.....S..V..A..LT.I.....S.KELE..PK---F.		
Artemia	SN..T.....S.....KE.C.....KTTEAEPK---FI		
Yeast	S.....S.R.D.LLE.N.....S.KKLEDHPK---F.		
Mucor	S..A.....S.LIE.I.....SKMEDSPK---FD		
ABP-50	KNGDAAMVELTSPSRMCMVESFTEYPLGRFAVRDMRQTAVGVKSTVKK	438	
Human	I..DMV.GK.....SD.....AVD..		
Tomato	G..KMI.TK..V..T.A.....V.NVD..		
Artemia	IT.V..K.L..A.SDE.....VNF.		
Yeast	L.KFV..K.....A.S.....VD.T		
Mucor	S..I..KMV..K.....AY.D.....AVE.V		
ABP-50	APGKAGDKKGAAPSKKK	456	
Human	..A.AGKVT..S.QKAQ..A.		
Tomato	D.TG.KVT..A.QKKG..		
Artemia	D.TAGKVT..A.EKAG..		
Yeast	E-KAAKVT..A.QKAA..		
Mucor	D-KAGKVT..A..KA...		

b

Depactin	(1)	* PQSGTALDENVKEIRAFKM (20)
		:
ABP-50	(164)	NYSQARYDEIVKEVSSFIKK (183)

one another for 98% of the entire coding sequence and define an identical polypeptide, differ slightly at the third position of redundant codons for some amino acids (Fig. 1).

To confirm that the cDNA sequences code for ABP-50, direct peptide sequencing of the purified protein was carried out. The N terminus was blocked, but sequencing of four internal tryptic peptides resulted in perfect matches with the amino-acid sequence deduced from the cDNA sequences (Fig. 1). In addition, the 2-nitro-5-thiocynobenzoic acid cleavage pattern and the amino-acid composition of ABP-50 were as predicted by the cDNA sequences (data not shown).

FIG. 1. The two cDNA sequences (p50-1.7, upper case; p50-1.5, lower case) of ABP-50 and the deduced amino-acid sequence (single-letter code). Bases are numbered from 5' to 3' in the left column and amino acids are numbered in the right column. The AATAAA poly(A)-addition signal, the sequenced tryptic peptides and the redundant codons (codons coding for identical amino acid) are underlined. Different codons used by the two cDNAs indicate the presence of two EF-1a genes as explained in the text. The λ gt11 cDNA library was screened with the antibodies made in our laboratory⁹. Isolated cDNA clones were subcloned into the *Eco*RI site of the plasmid (pBluescript II SK +, Stratagene) and partially deleted by successive *Exo*III exonuclease digestion (Erase-a-Base system, Promega). The nucleotide sequences were determined using a DNA sequencing kit (Sequenase 2.0, US Biochemical). All experiments were performed using standard procedures described elsewhere²², or according to the manufacturer's instructions. Purified ABP-50 was prepared as previously described⁹ and tryptic digestion was performed for 15 h in 2 M urea and 0.1 M ammonium bicarbonate. Peptides were resolved on a Vydac C18 HPLC column using an acetonitrile gradient. Selected tryptic peptides were sequenced in the Laboratory for Macromolecular Analysis at Albert Einstein College of Medicine and the Microchemistry Facility at the Biological Laboratory, Harvard University.

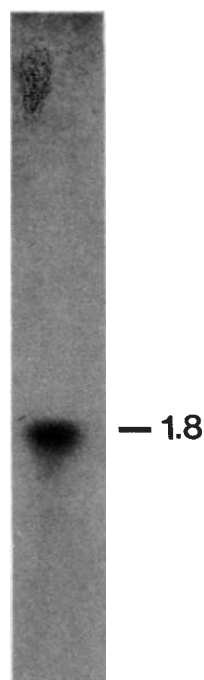


FIG. 3 Northern blot of RNA extracted from vegetative AX-3 cells and probed with ^{32}P -labelled p50-1.5. The probe was made with a labelling kit (Multiprime DNA Labelling System, Amersham) following the manufacturer's instructions. Procedures for RNA extraction, blotting and probing were performed as described by Franke *et al.*²⁴. The RNA size (kb) is indicated.

Sequence analysis demonstrates that ABP-50 shares 73–76% identity with EF-1a from human fibroblast (76%)¹¹, tomato (75%)¹², brine shrimp (74%)¹³, yeast (73%)¹⁴ and a fungus (73%)¹⁵ (Fig. 2a). The GTP and tRNA binding regions^{14–19} are conserved. Comparison of ABP-50 with other actin-binding proteins showed that a region of ABP-50 adjacent to the guanine nucleotide binding site is homologous to the actin-binding domain of depactin (Fig. 2a, b). Depactin is an actin-binding protein isolated from starfish oocytes. Protein crosslinking experiments have implicated residues 1–20 at the N terminus of depactin in actin binding²⁰. A weak homology was found in the putative depactin-like actin-binding motif among the EF-1a sequences from a broad range of species (Fig. 2a).

The composite cDNA sequence of 1.8 kb is consistent with northern blot analysis in which only one band at 1.8 kb is detected using p50-1.5 as the probe (Fig. 3). Size heterogeneity at the protein level can be ruled out because the polyclonal antibodies against ABP-50 detect a single 50K band in a whole cell lysate⁹. The distinction between the two cDNA sequences indicates the presence of at least two genes coding for EF-1a in *D. discoideum*. EF-1a encoded by multiple genes has been found in other eukaryotic cells, such as human fibroblasts, tomato, brine shrimp, yeast and a fungus^{11–15}. Southern analysis also suggests the presence of two, but no more than two, genes (data not shown), which nevertheless does not exclude the possibility of a single gene interrupted by intron(s) with several restriction sites. The identification of two cDNAs that differ in 34 redundant codons but still code for an identical polypeptide is highly unlikely to result from cloning and sequencing artefacts but is consistent with the presence of two genes. Similar observations of slight differences in the coding region of the cDNA clones, the genes, and between sequenced cDNAs and peptides have been documented in studies of EF-1a factors from tomato, brine shrimp and human fibroblasts. They are apparently the result of multiple genes^{11–13}.

Analysis of purified ABP-50 using a poly(U)-directed *in vitro* translation system²¹ demonstrates that ABP-50 is a functional EF-1a, as it catalyses polyphenylalanine synthesis in a concentration-dependent fashion indistinguishable from that of rabbit reticulocyte EF-1a (Fig. 4).

On the basis of the functional analysis, sequence comparison and the studies using antibody and cDNA probes that detect

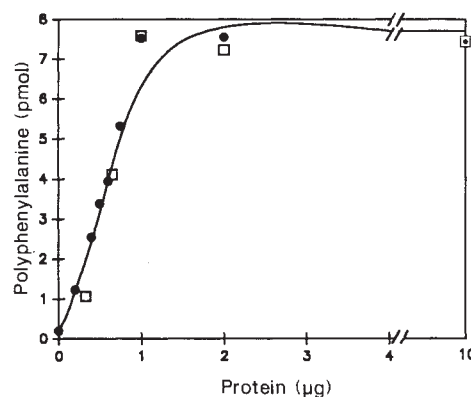


FIG. 4 Concentration dependence of polyphenylalanine synthesis by ABP-50. Catalysed polyphenylalanine synthesis (y-axis) in a poly(U)-directed *in vitro* translation system²¹ was assayed by adding varying amounts (x axis) of ABP-50 (●) or rabbit reticulocyte EF-1a (□) to the reaction mixture containing 0.4 A₂₈₀ units of purified rabbit reticulocyte ribosomes, 0.4 μg rabbit reticulocyte EF-2, 10 pmol [^3H]tRNA^{Phe} (Amersham), 4 μg poly(U), 10 mM MgCl₂, 1 mM GTP, 20 mM Tris (pH 7.5), 100 mM KCl, 2 mM phosphoenolpyruvate, and 1 mM DTT in 50 μl of reaction mixture. The reaction mixture was incubated for 10 min at room temperature and then assayed for pmol polyphenylalanine retained by a nitrocellulose filter. Rabbit reticulocyte ribosomes, EF-1a, and EF-2 were gifts from William Merrick of Case Western Reserve University. Other reagents were purchased from Sigma.

only one species in total cellular proteins and RNAs respectively, we conclude that ABP-50 is EF-1a of *D. discoideum*. The identity of EF-1a and ABP-50 provides the first direct evidence that an element of the protein synthetic apparatus binds to the actin cytoskeleton. EF-1a is an enzymatic factor catalysing the binding of the aminoacyl-tRNAs to the ribosomal particle in the peptide elongation phase of protein synthesis. It may therefore serve as a target for protein synthesis regulation. The binding of EF-1a to actin may affect the activity or availability of EF-1a in protein synthesis. Furthermore, the binding to either G-actin or F-actin may have different effects on EF-1a function. As the incorporation of EF-1a into the cytoskeleton is regulated during chemotaxis as shown in our previous studies¹⁰, the reversible association of EF-1a with actin may provide a mechanism for temporal and spatial regulation of protein synthesis in eukaryotic cells. □

Received 5 April; accepted 12 August 1990.

1. Lenk, R., Ransom, L., Kaufmann, Y. & Penman, S. *Cell* **10**, 67–78 (1977).
2. Nielsen, P., Goetz, S. & Trachsel, H. *Cell Biol. Int. Rept.* **7**, 245–254 (1983).
3. Bonneau, A., Darveau, A. & Sonenberg, N. *J. Cell Biol.* **100**, 1209–1218 (1985).
4. Ornelles, D. A., Fey, E. G. & Penman, S. *Molec. cell. Biol.* **6**, 1650–1662 (1986).
5. Cheng, H. & Bjerknes, M. *J. molec. Biol.* **210**, 541–549 (1989).
6. Singer, R. H., Langevin, G. L. & Lawrence, J. B. *J. Cell Biol.* **108**, 2343–2353 (1989).
7. Kuriyama, R., Savereide, P., Lefebvre, P. & Dasgupta, S. *J. Cell Sci.* **95**, 231–236 (1990).
8. Ohta, K. *et al. J. Biol. Chem.* **265**, 3240–3247 (1990).
9. Demma, M., Warren, V., Hook, R., Dharmawardhane, S. & Condeelis, J. *J. Biol. Chem.* **265**, 2286–2291 (1990).
10. Dharmawardhane, S., Demma, M., Warren, V. & Condeelis, J. *J. Cell Biol.* **109**, 273a (1989).
11. Uetsuki, T., Naito, A., Nagata, S. & Kaziro, Y. *J. Biol. Chem.* **264**, 5791–5798 (1989).
12. Pokalsky, A. R. *et al. Nucleic Acids Res.* **17**, 4661–4673 (1989).
13. van Hemert, F. J., Amons, R., Pluijms, W. J. M., van Ormondt, H. & Möller, W. *EMBO J.* **3**, 1109–1113 (1984).
14. Nagata, S. *et al. EMBO J.* **3**, 1825–1830 (1984).
15. Uetsuki, T., Lira, L. M. & Sypherd, P. S. *J. Biol. Chem.* **261**, 15022–15029 (1986).
16. Tabin, C. J. *et al. Nature* **300**, 143–149 (1982).
17. Jurnak, F. *Science* **230**, 32–36 (1985).
18. Taparowsky, E., Shimizu, K., Goldfarb, M. & Wigler, M. *Cell* **34**, 581–586 (1983).
19. Bosch, L. *et al. Trends biochem. Sci.* **10**, 313–316 (1985).
20. Sutoh, K. & Mabuchi, I. *Biochemistry* **28**, 102–106 (1989).
21. Merrick, W. C. *Meth. Enzym.* (eds Moldave, K. & Grossman, L.) **60**, 108–123 (Academic, New York, 1979).
22. Davis, L. G., Dibner, M. D. & Battey, J. F. *Basic Methods in Molecular Biology* (Elsevier, New York, 1986).
23. Woolley, P. & Clark, B. F. C. *Bio/technology* **7**, 913–920 (1989).
24. Franke, J., Podgorski, G. J. & Kessin, F. H. *Dev. Biol.* **124**, 504–511 (1987).

ACKNOWLEDGEMENTS. We thank R. Kessin, L. Leinwand and the members of their labs for introducing us to molecular biology, and J. Segall, S. Frankel, A. Bresnick, R. Eddy, D. Cox, and R. Sauterer for critical reading of the manuscript. This work was supported by the NIH.

Antibodies from *Escherichia coli*

A. Plückthun

Use of *Escherichia coli* as an expression host has opened up new possibilities in antibody research and its applications. It greatly facilitates rational engineering and random mutagenesis.

THE use of *E. coli* as an expression host brings the arsenal of bacterial gene technology to antibody molecules. The ease with which *E. coli* can be manipulated greatly facilitates antibody engineering of both the antigen binding pocket and the framework. The efficient transformation and transfection permit random mutagenesis experiments and the construction of libraries. The fast growth and easy fermentation of *E. coli* allow rapid, large-scale production of antibody fragments, which is a prerequisite for structural studies. The bacterial metabolism permits easy labelling of antibody protein with isotopes, and one may also devise metabolic selection schemes for binding and catalysis by an antibody.

The expression system developed by us^{1,2} and independently by Better *et al.*³ combines the advantages of antibody expression in the native state with the use of *E. coli* as the host. Expression in the native state is essential for any kind of screening as the individual refolding of many random mutants *in vitro* would not be feasible. Functional expression is also a prerequisite for the one-step purification process of antigen-affinity chromatography^{1,2,4}. Finally, different refolding properties of mutants may obscure areas of interest in studies of structure and function.

Expression system

A strategy was developed that consists of the simultaneous secretion of both chains of an antibody fragment (Fig. 1) into the periplasm of *E. coli*. The two chains can therefore fold in each other's presence, and were found to assemble correctly *in vivo*. The disulphide formation within each variable domain, which was found to be essential for stability⁵, can occur in the oxidizing milieu of the periplasm. It was noted that the outer membrane can become leaky, in which case the periplasmic content is lost to the medium², independent of the choice of signal sequence⁵. Lysis can be prevented, however, by growing the cells at lower temperatures. The final yield can vary depending on the antibody fragments (Fig. 1) used for expression^{1,2,4,5}.

Our studies were carried out with the anti-phosphorylcholine antibody McPC603^{6,7}. This antibody has been extensively characterized, including the determination of its crystal structure with and without bound antigen^{8,9}.

Antigen binding fragments of an antibody:

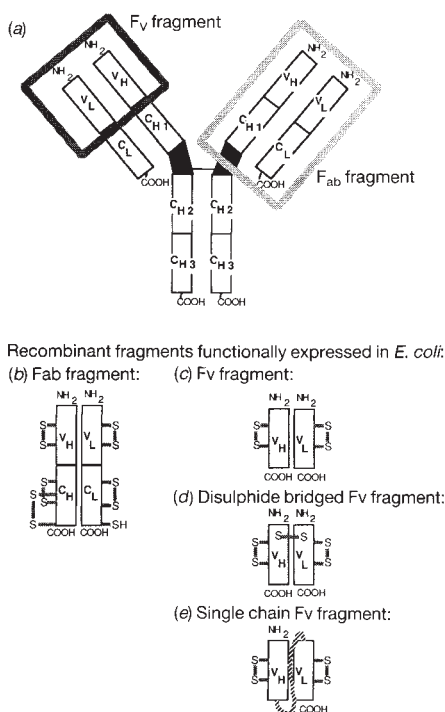


FIG. 1 Arrangement of the various antibody domains in the complete antibody and recombinant fragments functionally expressed in *Escherichia coli*. Note that the disulphide bond structure drawn in the Fab fragment is that of McPC603, a mouse IgA.

Recombinant Ab fragments

Fab fragments (Fig. 1 a and b) of antibodies, which can easily be prepared by proteolytic cleavage, are well documented as having the same antigen binding affinities as whole antibodies. The same applies to the recombinant Fab fragment of McPC603 from *E. coli*^{2,5,10}. Fab fragments are stable and associate well because the constant domains contribute to stability. However, the yield of the functional Fab fragment of McPC603 (ref. 5) is significantly lower than that of the Fv fragment (see below; Fig. 1 c) in *E. coli*. An analysis of this phenomenon with several diagnostic framework variants⁵ revealed that the problem does not lie in transport or processing, but rather in folding and/or assembly *in vivo*.

The Fv fragment (Fig. 1 c) is the smallest conceivable fragment that still contains the complete binding site. Fv fragments

are difficult, often impossible, to prepare by proteolysis, and little was known about their properties before they became available through gene technology¹. Analysis has shown^{1,4} that the Fv fragment does have the same binding properties as the Fab fragment or whole antibody. However, Fv fragments have a tendency to dissociate into V_H and V_L upon dilution¹. The exact equilibrium constants will vary from antibody to antibody, as the hyper-variable loops contribute to this interaction. For McPC603, the dissociation constant was found to be about 10⁻⁶ M (ref. 4). This problem does not interfere with all uses, and the relatively small size of the molecule makes it an attractive target for structural studies by X-ray crystallography and nuclear magnetic resonance. Fv fragments may have important applications in the diagnosis and treatment of tumours, where their small size may facilitate the penetration of dense tumour tissue and provide for low antigenicity.

Several strategies have been developed to overcome the dissociation problem while maintaining the small size⁴. First, the two chains can chemically be cross-linked with glutaraldehyde. With McPC603, this can be accomplished without any loss of binding affinity. Second, an intermolecular disulphide bond can be created (Fig. 1 d). It has been shown that the additional disulphide bridge forms spontaneously in the periplasm and the molecule can, therefore, be obtained in fully functional form from *E. coli*². Antigen binding affinity was only marginally affected. Third, the two domains can be connected by a peptide linker (Fig. 1 e). Such single chain Fv fragments had been obtained previously as insoluble inclusion bodies that had to be refolded *in vitro*^{11,12}. The single chain Fv fragment can also be secreted in functional form with almost unaltered binding affinity⁴. Linking the domains in the secreted single-chain fragment does not significantly influence the yield of functional fragment as compared to the secreted Fv fragment, and it is, therefore, unlikely that the association of V_L and V_H is a kinetic problem in *E. coli*. Consequently, the linking of the two domains is optional, and not essential. It remains to be seen whether the linker may obstruct other antigen binding sites: it does not in McPC603.

The properties of the isolated domains V_H and V_L were also investigated. Again,

their properties will vary depending upon the antibody under study. With McPC603, no binding of V_H to the antigen could be shown. The V_H domain also shows low solubility at temperatures above about 4 °C. Consequently, the use of V_H domains as a general substitute for antibody-combining sites¹³ presents some technical challenges. While the solubility problem may be alleviated by engineering the framework, it remains to be shown that satisfactory binding and, most importantly, narrow specificity can be maintained in such modified fragments.

On the other hand, V_L has a tendency to dimerize with itself. The precise dissociation constant again varies from antibody to antibody, but is usually between 10^{-3} M and 10^{-6} M. V_L of McPC603 does not bind the antigen either, but this recombinant domain does give rise to well ordered crystals of the dimer¹⁴. This offers the exciting prospect of rapidly obtaining a structural database of characteristic complementarity-determining regions.

Libraries

The expression of antibody libraries^{13,15} is an important extension of the expression of defined single species of antibodies¹⁴. The advent of polymerase chain reaction (PCR) technology has provided rapid access to the family of antibody genes as they contain conserved sequences in the 5' and 3' portion of variable genes, the framework region and the whole constant region. In one approach, a library of λ phages was generated, recombined from individual PCR products obtained from mRNA for antibody heavy and light chains isolated from an immunized animal¹⁵. In the recombinant phage, the genes were arranged exactly as in the expression vectors described previously¹⁻³. Fab fragments with a desired binding activity can be found by a 'reversed' immunoscreening with labelled antigen, although phage may not be such a good production vehicle.

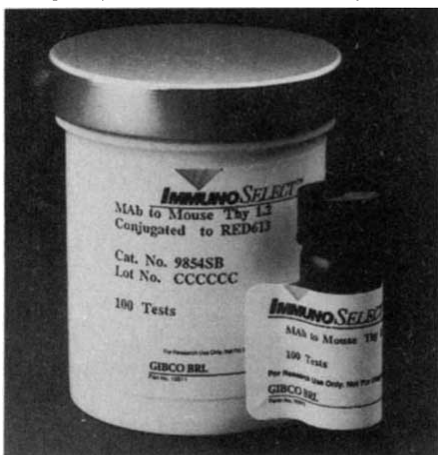
The real challenge with large antibody libraries, however, is screening. While conventional screening technology may be useful with smaller libraries, for example, from the highly enriched mRNA of an immunized animal, the task of finding a binding species with high affinity and narrow specificity out of the complete spectrum of conceivable antibodies (that is, of a non-immunized animal) is formidable indeed. Therefore, the potential substitution of monoclonal antibody technology by *E. coli* is likely to depend mainly on developing screening and enrichment systems for narrow binding specificities that measure up to the efficient multistep processes that the immune system uses.

Human antibodies are one of the obvious targets for this technology. Such human antibody fragments should signi-

Today's immunology

A miniblott system with disposable cassettes, monoclonal antibodies to dystrophin and a new fluorochrome for flow cytometry — new product news for the immunologist.

WHEN used in conjunction with fluorescein- and phycoerythrin-labelled antibodies, the new **fluorochrome** — RED613 — from Life Technologies allows researchers to perform simultaneous multicolour immunofluorescence analysis (*Reader Service No. 101*). When



RED613: fluorochrome for flow cytometry, used with an argon laser, the company says that RED613 produces a stronger signal at its emission peak than that of other dyes used in the labelling of monoclonal antibodies for flow cytometry. Monoclonal antibodies to mouse CD3, CD4, CD8a, CD25, CD45, CD45R, CD45R-1, H-2^{kbqst}, Ia and Thy 1.2 are available conjugated to RED613. These products are supplied in 25-, 100- and 200- μ g quantities.

The HIV LiaTek system from Organon Teknika is a strip-based **enzyme immunoassay for the differential detection**

of antibodies to HIV 1 and 2 in human serum or plasma (*Reader Service No. 102*). Specific recombinant proteins and synthetic peptides derived from HIV-1 and HIV-2 are coated as discrete lines onto nylon strips. Four HIV-1 antigens are applied: the core protein p24, an additional core protein p17, the gp41 region of the envelope gene and the p31 endonuclease domain of the viral polymerase gene. An antigen for the transmembrane glycoprotein gp36 is applied for the detection of HIV-2. The EIA procedure is as follows: strips are incubated with the test samples and any unbound material is washed away; a second incubation is carried out with enzyme conjugate; the strips are washed again and then incubated with chromogen. A positive result is denoted by the presence of brown bands where alkaline phosphatase-labelled antibody has bound. A semi-quantitative interpretation of the results is possible, as colour development is proportional to the amount of specific antibody present, says Organon Teknika. Results can be obtained within 2.5 hours. The LiaTek system costs £290 (UK) for 20 tests.

Assorted monoclonals

DYS-1 and DYS-2 are two monoclonal antibodies to the protein product of the **Duchenne muscular dystrophy** gene, dystrophin, which were developed by researchers at Newcastle General Hospital and are now available from Novocastra Laboratories (*Reader Service No. 103*). The mAbs can be used to identify aberrant protein localization patterns in

ificantly reduce the problem of antigenicity in the clinical use of antibodies. Yet, there is another hurdle. If, however, antibodies with high affinity and high selectivity cannot be found with *E. coli* screening, problems with crossreactivity against undesired targets in the body might lead to severe clinical consequences, especially if a low-affinity antibody is given in high doses.

There are, therefore, many challenges ahead in antibody research. □

Andreas Plückthun is group leader at the Genzentrum der Universität München, c/o Max-Planck-Institut für Biochemie, Am Klopferspitz, D-8033 Martinsried, FRG. For more information, fill in reader service number 100.

- Skerra, A. & Plückthun, A. *Science* **240**, 1038–1041 (1988).
- Plückthun, A. & Skerra, A. *Meth. Enzym.* **178**, 497–515 (1989).
- Better, M., Chang, C.P., Robinson, R.R. & Horwitz, A.H. *Science* **240**, 1041–1043 (1988).
- Glockshuber, R., Malia, M., Pfitzinger, I. & Plückthun, A. *Biochemistry* **29**, 1362–1367 (1990).
- Skerra, A. & Plückthun, A. (submitted).
- Perlmutter, R.M. et al. *Adv. Immun.* **35**, 1–37 (1984).
- Plückthun, A., Glockshuber, R., Pfitzinger, I., Skerra, A. & Stadlmüller, J. Cold Spring Harb. Symp. quant. Biol. **52**, 105–112 (1987).
- Segal, D.M., Padlan, E.A., Cohen, G.H., Rudikoff, S., Potter, M. & Davies, D.R. *Proc. Natn Acad. Sci. U.S.A.* **71**, 4298–4302 (1974).
- Satow, Y., Cohen, G.H., Padlan, E.A. & Davies, D.R. *J. molec. Biol.* **190**, 593–604 (1986).
- Skerra, A., Glockshuber, R. & Plückthun, A. (submitted).
- Bird, R.E. et al. *Science* **242**, 423–426 (1988).
- Huston, J.S. et al. *Proc. Natn Acad. Sci. U.S.A.* **85**, 5879–5883 (1988).
- Ward, E.S., Güssow, D., Griffiths, A.D., Jones, P.T. & Winter, G. *Nature* **341**, 544–546 (1989).
- Glockshuber, R., Steipe, B., Huber, R. & Plückthun, A. *J. molec. Biol.* **213**, 613–615 (1990).
- Huse, W.D. et al. *Science* **246**, 1275–1281 (1989).

frozen tissue sections, or the presence of a protein with abnormal size or abundance on western blots of muscle tissue. DYS-1 recognises an epitope in the central rod domain of the molecule and reacts with many dystrophin metabolites, including the characteristic 220,000 Da degradation product, which occurs in most patients with Becker muscular dystrophy — the milder allelic condition. DYS-2 recognises an epitope at the extreme carboxy terminus. Both mAbs are, says Novocastra, suitable for immunohistochemistry, western blotting and immunogold ultrastructural studies. Aliquots of 2.5 ml DYS-1 and DYS-2 cost £95 (UK).

BioGenex Laboratories is offering a monoclonal antibody — AE8 — for the **specific and qualitative localization of cytokeratin 13** in paraffin-embedded tissue sections (*Reader Service No. 104*). AE8 reacts strongly with the acidic 51K (No. 13) keratin, which is expressed mainly in the suprabasal cells of oesophageal, ventral tongue, buccal mucosal, exocervical and many other internal,



AE8 mAb for cytokeratin localization.

stratified squamous epithelia, says BioGenex. The monoclonal cytokeratin antibody AE8 is available as a ready-to-use reagent, costing \$135 (US), or it can be supplied in kit form with either the \$205 (US) HistoGen (PAP) or the StrAviGen high-performance or super-sensitive detection systems, which cost \$205 (US) and \$255 (US), respectively.

T Cell Sciences has introduced the CD-SET of **monoclonal antibodies to the human CD antigens** — CD2, CD3, CD4 and CD8 (*Reader Service No. 105*). The CD-SET of mAbs recognise specific cell surface structures, which have been assigned a CD designation by the International Workshop on Human Leucocyte Differentiation Antigens, says T Cell Sciences. The mAbs can be used for flow cytometry, immunohistology and immunoprecipitation procedures to investigate the role of T lymphocytes in



CD-SET — mAbs to human CD antigens

lymphoid malignancies, autoimmune diseases and immune system disorders. The mAbs, which are sold in quantities sufficient for 100 tests, can be supplied in a purified or FITC-conjugated form for \$240 (US) or \$355 (US), respectively.

In kit form

The **mouse monoclonal antibody isotyping kit** from Holland biotechnology is designed to identify all mouse immunoglobulin classes and subclasses in 5–30 minutes and with a detection sensitivity of less than $1 \mu\text{g ml}^{-1}$ (*Reader Service No. 106*). The assay is based on a sol particle detection system, which eliminates the need for radioactive tracers, enzymes and substrates. The test is carried out by adding culture supernatant or diluted ascites to a tube that contains the test strip onto which subclass-specific antibodies have been spotted. A second solution containing anti-mouse antibodies coupled to a coloured sol particle is added. A positive test results in the formation of two coloured spots on the test strip: one identifying the isotype of the antibody and the other indicating if it is kappa or lambda. The isotyping kit can be supplied with enough reagents and test strips for 10 or 20 assays.

Janssen Biochimica has a ready-to-use radial immunodiffusion (RID) kit for the **quantification of human IgG subclasses** in serum (*Reader Service No. 107*). The polyclonal anti-human IgG subclass reagents produce sharp precipitation rings with the RID technique, react specifically with human IgG subclass reagents, but do not react with allotype and iso-allotype determinants, says Janssen. The kit includes four RID plates in a 3×8 -well format, a lyophilized prediluted control and two lots of lyophilized prediluted standard sera.

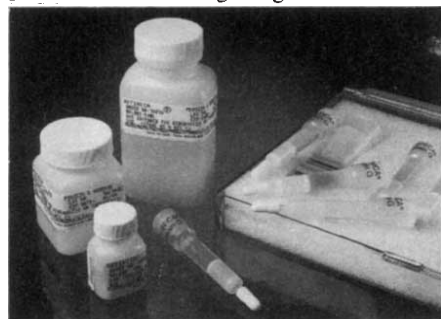
The IgM Affinisorb Kit from Windsor Park Laboratories provides everything you need to **isolate and purify mouse IgM monoclonal antibodies** from mouse ascites fluid in a one-step chromatographic procedure (*Reader Service No. 108*). The IgM affinity gel consists of an affinity ligand immobilized onto a cross-linked beaded agarose through a stable secondary amine linkage. The \$285 (US) kit includes binding and elution buffers and a 4 ml prepacked column of IgM affinity gel. One column has the capacity to purify 3–4 mg of mouse monoclonal IgM per run, says the company. Sufficient buffers are supplied to perform ten purification runs using the 4-ml column.

For researchers who perform standard immunological procedures such as immunocytochemistry, western blotting and *in situ* hybridization, British Biotechnology offers the **Light Up silver en-**

hancement kit (*Reader Service No. 109*). The £19 (UK) kit is designed to increase sensitivity when using gold-labelling reagents, and produce low background staining.

Pure and simple

Schleicher and Schuell has introduced Affinica Protein G agarose, which complements an existing range of Protein A



Protein G agarose from S & S.

ADVERTISEMENTS

TIB MOLBIOL BERLIN
CUSTOM DNA SYNTHESIS
5, — DM per base
DM 95,— per oligo
FAX (0049) 30 6261960

Reader Service No.79

Custom DNA
Purified and Shipped in 48 hours,
Worldwide.
100 ug, \$10.00 a base for the first 15 bases,
\$7.50 for each additional base.
Delivered.
Research Genetics
1-800-533-4363
In the United Kingdom 0-800-89-1393
In Canada 1-800-533-4363
FAX 205-536-9016

Reader Service No.26

Synthetic Oligonucleotides
Custom made Probes and Primers,
Synthesised, purified and shipped in 3 days
£4.00 per base
(All inclusive)
0.2 μM scale (approx. 500 μg)
Modified bases, aminoalkylation,
data base analysis and full design facilities
Molecular Medicine Unit,
King's College School of Medicine & Dentistry,
Denmark Hill, London SE5 8RX
Tel. XX-44-(0)71 - 326 3126
FAX XX-44-(0)71 - 733 3877

Reader Service No.56

ADVERTISEMENTS

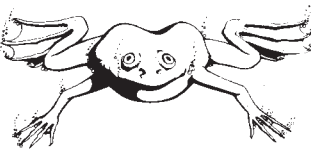
APP Δ QUALITY
 Δ INNOVATION
 Δ SERVICE

ADVANCED PROTEIN PRODUCTS LTD.
 UNIT 18H PREMIER PARTNERSHIP ESTATE
 BROOKMOOR, BRIEFLY HILL DY5 3UP
 Telex No. 338427 Fax No. 0384 480361
 Telephone No. 0384 263862

Δ FOETAL CALF SERUM — dialysed FCS available upon request.
 Δ APPRO-SERA — Immunoglobulin depleted bovine serum.
 Δ BOVINE SERUM ALBUMIN — only UK manufacturer.
 Δ TRANSFERRIN — bovine and porcine origin.
 Δ MONOCLONAL ANTIBODY RESOURCE SERVICE.

Reader Service No.20

SINGER
 INSTRUMENTS



**COMPLETE SYSTEMS FOR
 XENOPUS
 OOCYTE INJECTION**

INT. SALES: +44. 984. 40226.
Singer Instrument Co. Ltd.
 Roadwater, Watchet, Somerset,
 England, TA23 0QL. Tel: Washford (0984) 40226.
 Telex: 337492 Comcab G

Reader Service No.50

Serving now
BULK AMOUNTS
 of
**90 RESTRICTION - and
 20 MODIFYING ENZYMES**

We guarantee excellent quality
 and very competitive prices.
 Please inquire for quotations

AGS Angewandte Gentechnologie
 Systeme GmbH
 Rischerstraße 12, D-6900 Heidelberg 1
 West-Germany
 Tel.: Germany (6221) 81023
 Fax: (6221) 840610

Reader Service No.80

**CUSTOM-MADE & PURIFIED
 DNA PEPTIDES**

\$8/base
 50-100 µg
 40 hour delivery

Quotes available
 50-300 mg
 1-3 wks delivery

GUARANTEED QUALITY & QUANTITY

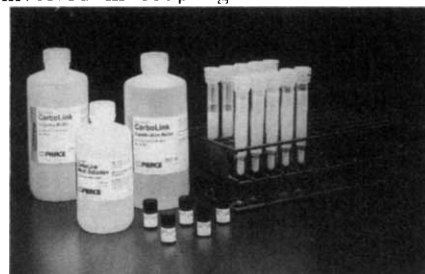
V.G. CORP.
 1200 Wonderland Rd. S
 Bldg #7, Unit #2
 London, Ontario
 Canada N6L 1A8

CALL OR FAX: (519) 652-3758

Reader Service No.29

and A/G **affinity purification gels** (*Reader Service No. 110*). The recombinant Protein-G ligand has two IgG binding domains and the albumin binding site from the native molecule has been eliminated. Protein G agarose binds only IgG — mouse, rat, cow, sheep, horse and goat, as well as human IgG₁ — but will not bind other immunoglobulins such as IgA, IgM or IgD, says S & S. Affinica Protein G agarose is available in prepacked 1-ml columns for \$50 (US), or in bulk 5-, 25- and 50-ml quantities.

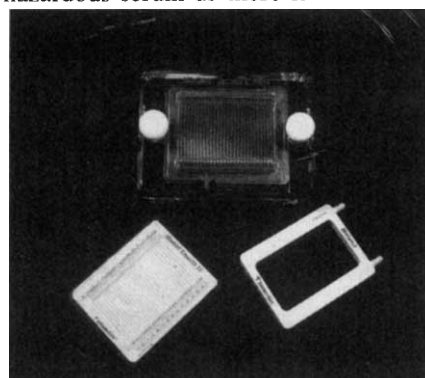
The ImmunoPure Ag/Ab Immobilization Kit #3 from Pierce is designed to provide a rapid and simple **method for orientating antibodies** (*Reader Service No. 111*). The kit features CarboLink coupling gel, which interacts with the IgG in the Fc region leaving both antigen binding sites free to interact with antigen in the mobile phase, says Pierce. The steps involved in coupling antibodies to the



Pierce's Ab/Ag immobilization kit.

CarboLink gel are: sugar groups are oxidized, allowing the *cis*-diols of the IgG to be transformed into reactive aldehyde moieties; these then combine with hydrazide groups on the CarboLink gel matrix to form a stable, leak-resistant linkage. The end result is, says Pierce, a reusable immunoaffinity column for anti-gen purification. The \$220 (US) kit includes five gel columns and sufficient reagents and buffers to prepare five 2-ml columns.

Screen up to 28 samples simultaneously on a single blot with the cassette Miniblot system for **western blot analysis** from Immunetics (*Reader Service No. 112*). The Miniblot system uses disposable cassettes; this prevents carryover from one assay to the next and minimizes contact with hazardous serum as there is no need to



Miniblot: the rapid screening system.

wash or decontaminate the incubation unit between runs, says Immunetics. Suitable applications for the system include the screening of antibodies for HIV and Lyme disease. A sealing system converts each channel of the cassette into a mini-incubation chamber, which eliminates cross-channel contamination, says Immunetics. As only 50 µl of antibody solution is required per test, the Miniblot system can be used to screen precious samples such as hybridoma supernatants and dried blood-spot eluates. Blots are washed between incubations by attaching a snap-on manifold to the cassette. The \$495 (US) Miniblot system consists of the cassette shell, Miniwash manifold and ten cassettes.

One-stop shopping

Bratton Biotech offers a range of **custom services** that may be of interest to the immunologist (*Reader Service No. 113*). BBT's services include: antibody production, ELISAs, blotting techniques, MHC typing, cloning, mutagenicity assays, flow cytometry, and applications of polymerase chain reaction technology.

The research and development services offered by Cymbus Bioscience now include the preparation of **non-rodent monoclonal antibodies** (*Reader Service No. 114*). Cymbus generates non-rodent mAbs — sheep, cattle and primate species — using heterohybridoma monoclonal antibody technology. Lymphocytes isolated from immunized animals are fused with heteromyeloma cell lines to generate stable hybrid clones secreting non-murine mAbs. This technology has already found therapeutic and prophylactic applications in veterinary medicine says Cymbus, because of the reduced anaphylactic potential of non-rodent mAbs as compared to mouse mAbs. Other areas of potential include acute toxin elimination, direct blockading of viral entry, or the specific targeting of drugs.

Contract conjugations are now available when using the antibody purification/conjugation service that Serotec has introduced (*Reader Service No. 115*). Antibodies or antibody fragments can be supplied conjugated to phycoerythrin, horseradish peroxidase and fluorescein isothiocyanate. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. To obtain further details, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

CORRECTION

The 5 July issue of *Nature* incorrectly stated that the DX-100 integrated, single-channel ion chromatograph was manufactured by Waters. It should have read Dionex.